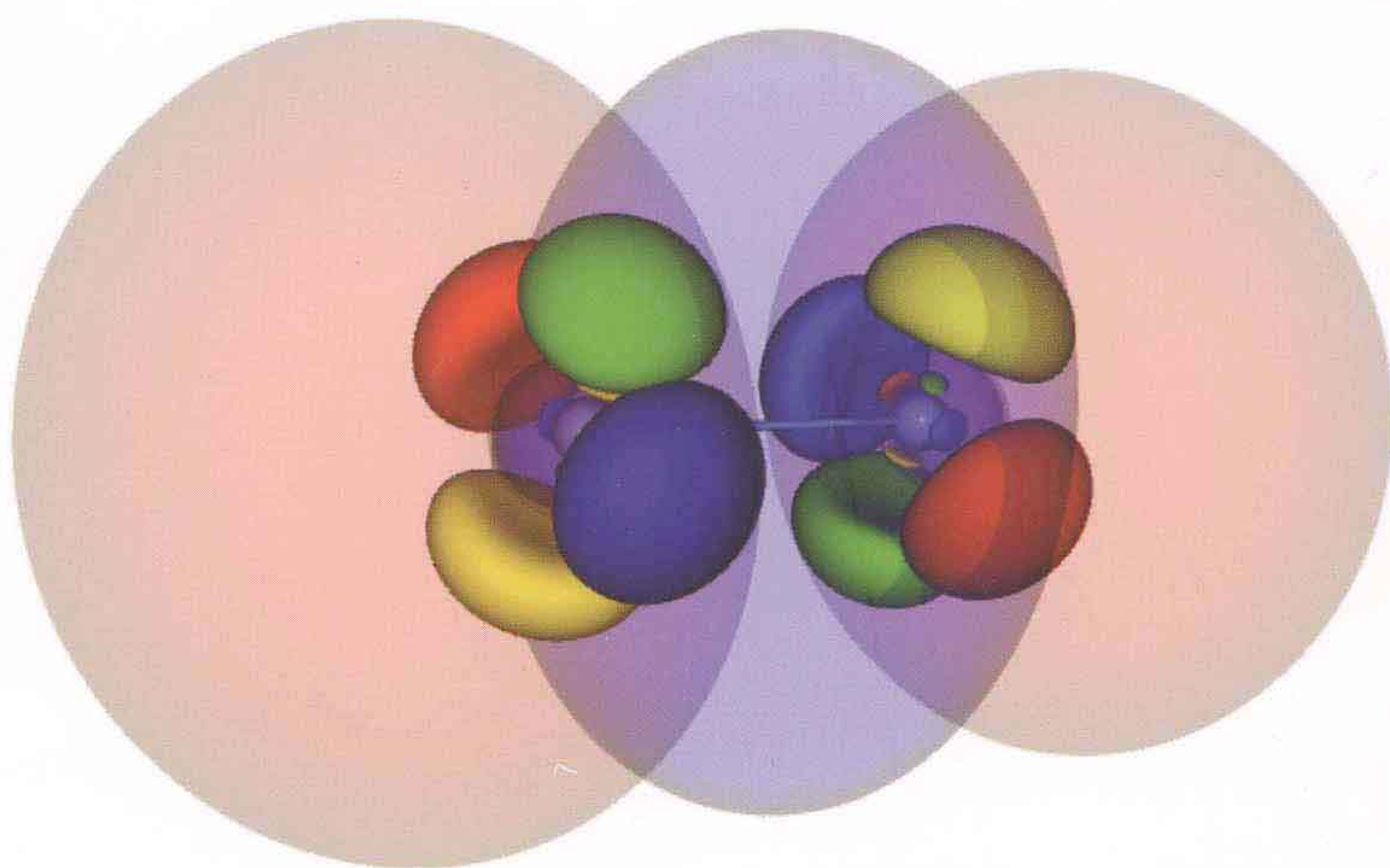
现代化学基础丛书 33

相对论量子化学基础

曹晓燕　方维海　〔德〕M.多尔格 著



双原子分子Tl_2的成键轨道σ、反键轨道π的相对论旋量表示

 科学出版社

(O-5182.0101)

现代化学基础丛书 33 相对论量子化学基础

- 详细推导了广泛用于相对论量子化学计算的几个二分量近似哈密顿量，包括Douglas-Kroll-Hess（DKH）哈密顿量、Wood-Boring（WB）哈密顿量（或称Cowan-Griffin哈密顿量）、Chang-Pelissier-Durand哈密顿量（或称zeroth-order regular approximation（ZORA）哈密顿量）、Pauli哈密顿量等。
- 在复杂体系的相对论量子化学计算中，有效核实势法是目前最流行的方法。其分支包括能量一致赝势、形状一致赝势、从头算模型势等，这些内容在本书中均有详细介绍。
- 按元素周期表中主族元素、过渡金属元素、镧系和锕系元素、超重元素的顺序，讨论了相对论量子化学理论在原子、分子电子结构计算中的应用。所总结的规律对实验工作者也是很有参考价值的。

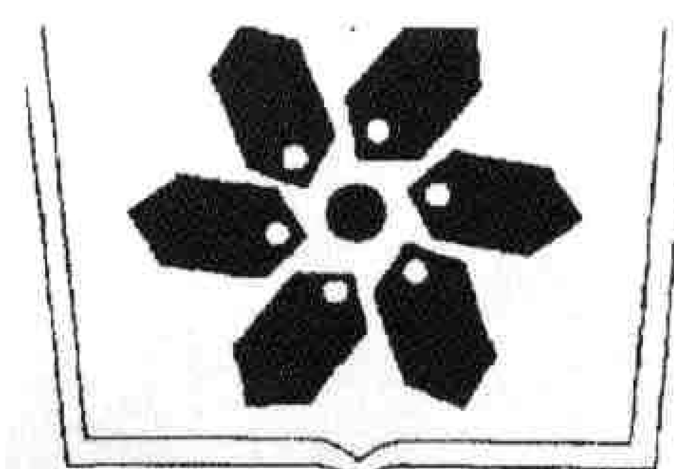
www.sciencep.com
ISBN 978-7-03-037959-7



科学出版社·科学化工分社
电 话：010-64001053
E-mail:sciencepress@foxmail.com

销售分类建议：化学/物理化学

定 价：60.00 元



中国科学院科学出版基金资助出版

现代化学基础丛书 33

相对论量子化学基础



曹晓燕　方维海　〔德〕M. 多尔格　著

科 学 出 版 社

北　京

内 容 简 介

本书是作者根据多年在德国科隆大学化学系教学与科研工作的经验而写成，主要介绍相对论量子化学的理论基础及计算方法。理论基础部分涵盖Einstein的狭义相对论、相对论波动方程、相对论的对称群、相对论的电子–电子相互作用等。相对论的计算方法包括全电子法和价电子法。在全电子法中，除四分量方法外，将重点介绍以下几个二分量近似 Hamilton 量，即Douglas-Kroll-Hess (DKH) Hamilton 量、Wood-Boring (WB) Hamilton 量、零级规则近似(zeroth-order regular approximation，ZORA) Hamilton 量、Pauli Hamilton 量等。在价电子法（又称有效核实势法）中，本书重点介绍能量一致赝势、形状一致赝势、从头算模型势等方法。

本书适宜作为量子化学专业研究生的教学参考书，也是相对论量子化学软件使用者的一本有用的参考书。

图书在版编目(CIP)数据

相对论量子化学基础/曹晓燕，方维海，(德)多尔格(Dolg, M.)著. —北京：科学出版社，2013

(现代化学基础丛书 33/朱清时主编)

ISBN 978-7-03-037959-7

Ⅰ. ①相… Ⅱ. ①曹… ②方… ③多… Ⅲ. ①相对论 ②量子化学 Ⅳ. ①O412.1 ②O641.12

中国版本图书馆 CIP 数据核字 (2013) 第 135976 号

责任编辑：周巧龙 孙 艳／责任校对：宣 慧
责任印制：钱玉芬／封面设计：陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2013 年 6 月第 一 版 开本：B5(720×1000)
2013 年 6 月第一次印刷 印张：13 1/4
字数：262 000

定价：60.00 元

(如有印装质量问题，我社负责调换)

《现代化学基础丛书》编委会

《现代化学基础丛书》序

如果把牛顿发表“自然哲学的数学原理”的 1687 年作为近代科学的诞生日，仅 300 多年中，知识以正反馈效应快速增长：知识产生更多的知识，力量导致更大的力量。特别是 20 世纪的科学技术对自然界的改造特别强劲，发展的速度空前迅速。

在科学技术的各个领域中，化学与人类的日常生活关系最为密切，对人类社会的发展产生的影响也特别巨大。从合成 DDT 开始的化学农药和从合成氨开始的化学肥料，把农业生产推到了前所未有的高度，以致人们把 21 世纪称为“化学农业时代”。不断发明出的种类繁多的化学材料极大地改善了人类的生活，使材料科学成为了 20 世纪的一个主流科技领域。化学家们对在分子层次上的物质结构和“态–态化学”、单分子化学等基元化学过程的认识也随着可利用的技术工具的迅速增多而快速深入。

也应看到，化学虽然创造了大量人类需要的新物质，但是在许多场合中却未有效地利用资源，而且产生了大量排放物造成严重的环境污染。以至于目前有不少人把化学化工与环境污染联系在一起。

在 21 世纪开始之时，化学正在两个方向上迅速发展。一是在 20 世纪迅速发展的惯性驱动下继续沿各个有强大生命力的方向发展，二是全方位的“绿色化”，即使整个化学从“粗放型”向“集约型”转变，既满足人们的需求，又维持生态平衡和保护环境。

为了在一定程度上帮助读者熟悉现代化学一些重要领域的现状，科学出版社组织编辑出版了这套《现代化学基础丛书》。丛书以无机化学、分析化学、物理化学、有机化学和高分子化学五个二级学科为主，介绍这些学科领域目前发展的重点和热点，并兼顾学科覆盖的全面性。丛书计划为有关的科技人员、教育工作者和高等院校研究生、高年级学生提供一套较高水平的读物，希望能为化学在新世纪的发展起积极的推动作用。

朱清时

序

相对论量子力学的奠基人 Dirac 在 1929 年发表的一篇论文中认为:“在考虑原子、分子结构和普通化学反应时，相对论效应是不重要的，忽略电子质量随速度的相对论变化，并且假定电子间及其与原子核间的作用只是 Coulomb 力，一般是足够精确的。”这是考虑到原子、分子中电子的平均运动速度远比光的速度小作出的推论。他的观点被普遍接受长达 40 年，早期量子化学计算中都忽略相对论效应。20 世纪 70 年代以后，人们逐步认识到，相对论效应在很多情况下起重要作用，不能忽略；否则很多实验事实，特别是涉及重元素的化学行为，不能得到合理解释。于是相对论量子化学随之发展起来，成为活跃的研究领域，促进了理论化学乃至整个化学学科的发展。现在对化学和物理过程的相对论效应进行计算模拟已经很常见了。由于 Dirac 方程的解是四分量的，存在对应于正电子的负能态解，在相对论量子化学计算中存在一些在近似求解 Schrödinger 方程时不存在的问题。例如，变分法求对应于电子的正能态解时可能出现变分塌陷，电子相关能的变分计算会遇到 Brown-Ravenhall 病等，需要用到专门的处理办法。可能有部分研究生和年轻化学工作者，由于缺乏相对论量子化学的基础知识，不能很好地理解与正确使用相关程序和解释计算结果。尽管介绍相对论量子化学的外文书籍比较多，但中文的相关书籍很少。为方便国内读者，出版一本介绍相对论量子化学基础知识的中文图书是很有必要的。

《相对论量子化学基础》一书作者中的 Dolg 教授是相对论量子化学领域的著名学者。他和曹晓燕博士长期从事相对论量子化学的理论和计算模拟研究，活跃在该领域发展前沿，并为德国的量子化学研究生开设相关课程。2010 年夏，曹晓燕博士在北京师范大学做访问教授期间，为全国计算和理论化学暑期学校讲授了这一课程。该书是在她授课讲稿的基础上增减一些内容编写而成的。从狭义相对论出发，导出相对论波动方程和电子间相互作用的相对论表达式，介绍了在相对论框架下对称性原理的表述方式。由于实际体系的相对论波动方程只能用近似方法求解，书中详细介绍了目前广泛使用的多种近似计算方法，比较它们的优缺点，并重点介绍作者有深入研究和重要贡献的价电子赝势方法。就我所知，这是第一本系统介绍相对论量子化学基础的中文专门教材。该书内容涵盖了相对论量子化学的各个重要方面；原理阐述清楚，简明扼要；主要公式推导从头开始，演算过程详细；可以让读者对重要概念和公式的来龙去脉有清晰的认识。全书文字流畅，叙述深入浅出，很适合数学和物理基础比较薄弱的化学专业学生阅读。所附参考文献则方便作深

入研究的读者了解有关问题研究前沿的概况。

该书不但可以帮助读者掌握相对论量子化学的原理和发展现状的概貌，还将有助于广大使用相对论量子化学计算软件的用户了解相关的理论基础，提高对计算结果的分析能力。希望这本书的出版有助于促进我国相对论理论化学研究的进步和发展。

黎乐民

2013 年 1 月

前　言

1926 年，也就是 Schrödinger 提出非相对论 Schrödinger 方程的同一年，Klein 和 Gordon 分别独立发表了相同的、也是第一个相对论波动方程，为铭记两人的贡献，该方程被命名为 Klein-Gordon 方程。后来，人们在 Schrödinger 未发表的手稿中也发现了同样的方程，所以 Klein-Gordon 方程又称为相对论的 Schrödinger 方程，描述自旋为零的粒子。两年后，Dirac 发表了描述自旋为 1/2 的微观粒子运动的 Dirac 方程。

从时间上看，相对论和非相对论的波动方程几乎是同时建立的，但在很长一段时间里，科学家更多地关心如何精确求解非相对论的波动方程，相对论效应则被认为是可以忽略的，对相对论波动方程的处理也十分繁琐，更谈不上精确。至 20 世纪 70 年代，非相对论量子化学方法已日益成熟，包含周期表中下部化学元素的量子化学计算成为可能。由于这些元素有十分明显的相对论效应，非相对论计算的结果始终不能令人满意，人们这才意识到相对论效应的重要性。同时，电子计算机的发展也为相对论计算提供了可能。1963 年，Herman 和 Skillman[1] 计算了周期表中所有原子的轨道能量，估计了相对论效应的影响。几年后，Fricke 等[2] 将结果扩展到 172 号元素，他们的研究表明，对重元素，原子的结构变化非常大。对 1~120 号元素，Desclaux[3] 计算了原子轨道的相对论校正。1974 年，Rosén 和 Ellis[4] 首先发表了 XeF_2 分子的相对论 Dirac HF 计算。几乎是同时，Desclaux 和 Pyykkö[5] 计算了碳族元素氢化物 (MH_4, M=C, Si, Ge, Sn, Pb) 的结构，考虑相对论效应后他们发现 M—H(M=Ge, Sn, Pb) 的键长分别缩短了 0.6%, 2.3%, 5.6%。从此，在相对论量子化学领域每年都有一定数量的文章发表，近年来已达到每年上千篇。经过这几十年的发展，相对论量子化学计算已经成为计算化学一个新的、迅速成长的分支。

相对论效应真的可以忽略吗？回答当然是否定的。早在现代的相对论量子化学计算方法出现以前，科学家就发现传统的非相对论理论并不能解释某些实验现象。以第 11 族元素 (铜族元素)，即铜 (Cu)、银 (Ag) 和金 (Au) 为例，金总是表现出与其铜族元素身份不相符合的性质，如为什么铜族元素中只有金具有稳定的三价氧化态？金为什么呈现金黄色？非相对论的量子化学理论并不能解释这一现象[6]。图 1 给出了 MH 和 M_2 (M=Cu, Ag, Au) 的键长变化趋势。根据非相对论的计算结果，Cu—H 键长 $<$ Ag—H 键长 $<$ Au—H 键长。实验结果却显示 Au—H 键长 $<$ Ag—H 键长。同样的问题对 M_2 (M=Cu, Ag, Au) 也存在，即非相对论计算结果显示 Cu—Cu 键长 $<$ Ag—Ag 键长 $<$ Au—Au 键长，而实验结果表明 Au —Au 键长

比 Ag—Ag 键长稍短。从图 1 可以清楚地看到，只有相对论的计算结果才能得到第 11 族元素正确的键长变化趋势。实际上，相对论效应在决定金的化学和物理性质中起着很大的作用[6]。另一个相对论效应明显的例子是关于液铅电池的电能，最新的理论计算表明大约 81% 的电能来源于相对论效应的贡献。换句话说，如果没有相对论效应，汽车将无法启动[7]。

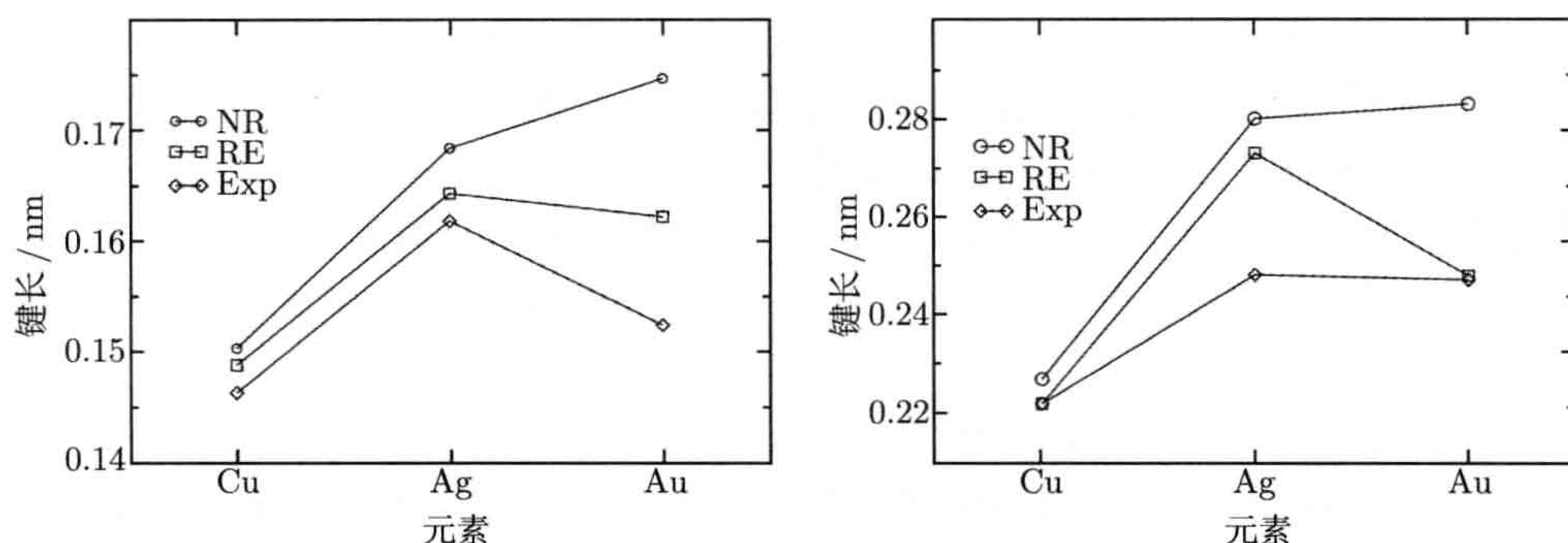


图 1　第 11 族元素双原子分子键长变化趋势图。左图为 MH 分子，右图为 M_2 分子 (M=Cu，Ag，Au)。NR：非相对论的计算结果; RE：相对论的计算结果; Exp：实验值

关于相对论量子化学研究进展的文献发表于各种科学杂志、会议报告和综述书籍等。有关相对论量子化学基本原理的综述文章有 Pyykkö[8, 9] 分别于 1978 年、1988 年发表在 *Adv. Quant. Chem.* 以及 *Chem. Rev.* 中的综述；Schwarz[10] 于 2010 年在 *Relativistic Methods for Chemists* 一书中撰写的第一章等。还有一些由多人合著的书籍综述了该领域各个研究方向的最新进展，如 Hess[11], Schwerdtfeger[12, 13], Hirao 和 Ishikawa[14], Barysz 和 Ishikawa[15] 等，由于这些综述书籍涉及的作者很多，以上所列人名均为编者姓名。此外，Pyykkö[16−18] 收集出版了 1916~1999 年发表的所有相关文献目录，为方便读者检索他还建立了关于原子和分子的相对论理论文献的网上数据库 (网址：http://rtam.csc.fi)。

就以中文为写作语言的发表物来看，刘文剑[19] 于 2007 年在《化学进展》发表了“相对论量子化学新进展”一文，详述了新一代准相对论方法，即精确的准相对论矩阵理论 (exact matrix quasi-relativistic theory，XQR)。徐光宪、黎乐民和王德民[20] 三人编撰的《量子化学 —— 基本原理和从头计算法》一书中部分章节也涉及相对论量子化学。

专门介绍相对论电子结构基础的英文书籍有三种，分别由牛津大学出版社 (Oxford University Press), Wiley-VCH, Springer 在 2007 年、2009 年、2010 年出版。第一本是 Dyall 和 Faegri[21] 合著的 *Introduction to Relativistic Quantum Chemistry*；第二本是 Reiher 和 Wolf[22] 合著的 *Relativistic Quantum Chemistry: The Fundamen-*

tal Theory of Molecular Science；第三本是 Grant[23] 编写的 *Relativistic Quantum Theory of Atoms and Molecules: Theory and Computations*。

本书的目的是给量子化学专业的研究生提供一本相对论量子化学的入门书籍，全书共分为七章：

第 1 章　狭义相对论基础　说明 Einstein 的两个假设以及由此带来的相对论效应，如传统的两个惯性参考系之间的变换公式 —— Galileo 变换公式仅在粒子低速运动时正确，当粒子以接近光速的速度运动时，则必须以 Lorentz 变换代替；粒子的质量和物体的长度将不再是常数，而是和选取的参考系有关；等等。

第 2 章　相对论波动方程　首先介绍第一个相对论波动方程，即描述自旋为零粒子运动的 Klein-Gordon 方程。重点介绍描述电子运动的 Dirac 方程 ——我们详细推导了自由电子、氢原子的 Dirac 方程的求解过程，让读者了解相对论四个量子数 (n, k, j, m_j) 或 (n, l, j, m_j) 的来历、为什么相对论波函数包含四个分量、何谓大分量 (小分量) 波函数等。通过与非相对论波函数的比较了解相对论波函数的一般特点，此章对学生建立相对论量子化学的基本认识非常重要。

第 3 章　相对论的电磁相互作用　在非相对论级别，光速被视作无穷大，两个带电粒子的相互作用是大家熟悉的 Coulomb 相互作用形式。相对论的电磁相互作用的严格推导需要用到量子电动力学的知识，这已超出了大部分化学专业的读者的知识范围。本章将在经典物理学的框架内，引入光速的有限性，导出低速条件下相对论的带电粒子相互作用，并以此和量子电动力学的结果相比较，最后得到目前在化学研究中最精确 (精确到 α^2 数量级①) 的 Dirac-Coulomb-Breit (DCB) Hamilton 算符。

第 4 章　相对论的对称性原理　由总角动量为半整数的量子态在绕任意轴的 360° 转动变换下改变符号这一性质，引导学生将非相对论的点群理论扩展到相对论的二倍群。重点介绍两个无限连续 Lie 群，即二维特殊酉群 SU(2) 和三维特殊实正交群 SO(3)。在徐光宪、黎乐民和王德民[20] 编撰的《量子化学 —— 基本原理和从头计算法》下册中有关于 Lie 群 Lie 代数的介绍。

第 5 章　全电子计算法　对大多数体系，我们必须用近似法来求解 Dirac 方程，本章介绍对体系全部电子作相对论量子化学计算的方法，即全电子方法。原则上，已知的计算方法如 Hartree-Fock (HF) 自洽场计算、微扰理论 (perturbation theory，PT)、组态相互作用 (configuration interaction，CI)、耦合簇 (coupled cluster，CC) 方法等均可以用于求解相对论的波动方程。由于相对论四分量波函数的特殊性，已有的非相对论计算程序并不能直接用于相对论计算。本章将介绍相对论 HF 方法和电子相关计算的一般原理，如非相对论能量的极小变分原理将由极

① α 为精细结构常数，约为 1/137

小极大变分原理取代，即大分量波函数变分取能量极小，小分量波函数变分取能量极大。由于大分量、小分量波函数的宇称相反，所以必须选择不同的基函数来展开这两组波函数，且二者基函数的选择必须满足基函数动能平衡条件以防止变分塌陷。为降低四分量计算带来的庞大计算量，人们也开发了所谓的二分量方法，一种是仍保留小分量，但分离自旋，在变分计算阶段不考虑自旋，而将自旋作微扰处理或在作 CI 计算时再加进自旋；另一种二分量方法是忽略小分量波函数对电子解的贡献，只考虑大分量波函数。为达此目的，一般是对分子 Hamilton 算符实施酉变换，如 Foldy-Wouthuysen 变换，简称为 FW 变换；Douglas-Kroll 变换，简称为 DK 变换；等等，变换后的 Hamilton 算符中导致大小分量耦合的奇算符的量级可降至 α^3 或更低，此时忽略奇算符，得到仅含偶算符的二分量 Hamilton 算符；另一种作法是直接忽略小分量波函数，求解仅含大分量波函数的 Hamilton 算符。目前使用广泛的二分量 Hamilton 算符有 Pauli(或 Breit-Pauli，BP) Hamilton 算符，Douglas-Kroll-Hess (DKH) Hamilton 算符，Wood-Boring (WB) Hamilton 算符，零级规则近似 (zeroth-order regular approximation，ZORA) Hamilton 算符等。本章将详细介绍如何经过推导得到以上 Hamilton 算符，并比较分析它们的优缺点。如果将自旋从二分量 Hamilton 算符中分离，则得到一分量 Hamilton 算符。第三种二分量方法是先引入基组，得到 Dirac 方程的矩阵表示，再通过变换消除小分量得到二分量的矩阵方程，此方法又称为“精确的二分量”(exact two-component，X2C) 方法。

第 6 章　有效核实势法　一般来说，相对论效应的大小大约和核电荷数的四次方成正比，所以在重元素化学研究中必须考虑相对论效应。另外，随着电子数目的增多，全电子计算量越来越大。受计算条件的限制，人们往往只能对原子和小分子作严格的全电子相对论电子结构计算。有效核实势 (effective core potentials，ECPs) 法的主要物理思想是将原子的内层电子连同原子核构成一个核实 (core)，该核实的电荷为原子的原子序数 (Z) 减去内层电子数 (n_c)。此时核实对外层价电子 (i) 的作用 (包括相对论效应的贡献) 以某种合适的有效核实势来表达。由于主要的相对论效应已包括在 ECPs 中，对外层价电子我们仍可以采用非相对论的 Hamilton 算符形式，进而进行变分和自洽迭代方法计算。ECPs 方法不仅能降低计算量，高精确度的 ECPs 还可以将高级别的相对论效应 (如完整的 Breit 相互作用) 隐含在 ECPs 中，而目前分子的全电子计算通常只包含 Breit 效应中的 Gaunt 相互作用。ECPs 方法有两个重要分支 —— 模型势 (model potentials，MPs) 和赝势 (pseudopotentials，PPs)，其中赝势也称为伪势。前者可以看成是对核实冻结条件下真实的价电子 HF 算符的模拟，所得到的原子的价电子轨道和真实的全电子轨道的径向节点 (即函数值为零的点) 结构是一样的。后者则是对全电子的价电子轨道进行变换，所得新轨道的径向节点数目一般比真实的价电子轨道的径向节点数目少。由于计算所需的基组大小与节点数目成正比，所以赝势的配套基组通常要比

模型势的配套基组更小一些。根据优化方法的不同，MPs 又细分为从头算模型势 (*ab initio* model potentials，AIMPs) 和模型核实势 (model core potentials，MCPs) 两种；赝势也有能量一致 (energy-consistent) 赝势及形状一致 (shape-consistent) 赝势两个分支。本章将详述这些不同 ECPs 的理论基础及如何优化得到它们的参数。

第 7 章　元素周期表中的相对论效应　本章讲述相对论量子化学理论在原子、分子电子结构计算中的应用。由于相关研究非常多，我们将按照主族元素、过渡金属元素、镧系和锕系元素、超重元素的顺序，分别选择有代表性的例子向读者介绍。

自 2007 年起，曹晓燕一直在德国科隆大学化学系讲授研究生课程《相对论量子化学基础》。她编写的课程讲义 (英文) 是本书的第一稿，其内容编排主要参考了① Pyykkö 编写的相对论量子化学讲义；② Dyall 和 Faegri[21] 合著的 *Introduction to Relativistic Quantum Chemistry*；③ 徐光宪、黎乐民和王德民[20] 编撰的《量子化学 —— 基本原理和从头计算法》。2010 年夏，曹晓燕在北京师范大学化学学院做访问教授期间，应邀在由中国教育部、国家自然科学基金委员会、北京师范大学三家单位共同举办的 2010 全国理论及计算化学暑期学校讲授了同名课程。由于这个机缘，曹晓燕、北京师范大学化学学院院长方维海教授、科隆大学理论化学研究所所长 Dolg 教授决定合作出版《相对论量子化学基础》一书。在曹晓燕英文讲义的基础上，Dolg 增删了部分内容，曹晓燕和方维海则负责中文稿的写作和定稿。我们希望呈现给大家的是一本易读、易懂的研究生读物。

与其他相对论量子化学的入门书籍[21, 22] 相比，本书注重基础知识的介绍，理论推导过程力求详细，更适合相对论量子化学初学者阅读。此外，本书作者长期从事有效核实势法研究，积累了丰富的经验，所以我们花了更多的篇幅详细介绍有效核实势法，这也是本书与其他书籍的不同之处。实际上，受计算条件的限制，在含重元素大分子体系的相对论量子化学计算中，人们更倾向于使用对计算机资源需求相对较小的有效核实势法。

感谢科隆大学理论化学研究所的博士生张鋆、美国普林斯顿大学杨军博士，他们详细阅读了本书初稿，发现了很多错误并提出了宝贵意见。感谢科隆大学理论化学研究所的 Börsch-Pulm 女士、Hanrath 博士和 Böhler 先生在中文写作环境的实现、绘图软件的使用及书稿打印等方面提供的帮助。

感谢芬兰赫尔辛基大学 Pyykkö 教授向中国科学院科学出版基金推荐本书。

特别感谢北京大学化学与分子工程学院的黎乐民院士，黎先生是我国稀土化合物理论研究 (包括相对论效应) 的先驱，他也是我国第一个相对论密度泛函程序 ——Beijing density functional (BDF) 的主要作者之一，非常感谢他向中国科学院科学出版基金推荐本书并为本书作序。

相对论量子化学的核心问题是相对论波动方程的求解，涉及非常广泛的数理

知识。作为一门新兴学科，相对论量子化学的研究正处在蓬勃发展的阶段。文献中各种方法、理论层出不穷，形成百花齐放的局面。本书尽可能地将目前流行的解决方案收罗其中，限于我们的水平，书中难免有不妥和疏漏之处，恳请读者批评指正。来函请发至 x.cao@uni-koeln.de，m.dolg@uni-koeln.de 或 fangwh@bnu.edu.cn。

参考文献

[1] Herman F, Skillman S. Atomic Structure Calculations. New Jersey Englewood Cliffs: Prentice-Hall Inc, 1963.

[2] Fricke B, Greiner W, Waber J T. The continuation of the periodic table up to $Z = 172$. The chemistry of superheavy elements. Theor Chim Acta, 1971, 21: 235~260.

[3] Desclaux J P. Relativistic Dirac-Fock expectation values for atoms with $Z = 1$ to $Z = 120$. At Data Nucl Data Tables, 1973, 12: 311~406.

[4] Rosén A, Ellis D E. Relativistic molecular wavefunctions XeF_2. Chem Phys Lett, 1974, 27: 595~599.

[5] Desclaux J P, Pyykkö P. Relativistic and non-relativistic Hartree-Fock one-center expansion calculations for the series CH_4 to PbH_4 within the spherical approximation. Chem Phys Lett, 1974, 29: 534~539.

[6] Schwerdtfeger P. Relativistic effects in gold chemistry. 2. the stability of complex halides of gold(III). J Am Chem Soc, 1989, 111: 7261~7262.

[7] Ahuja R, Blomqvist A, Larsson P, et al. Relativity and the lead-acid battery. Phys Rev Lett, 2011, 106: 018301-1~018301-4.

[8] Pyykkö P. Relativistic quantum chemistry. Adv Quant Chem, 1978, 11: 353~409.

[9] Pyykkö P. Relativistic effects in structural chemistry. Chem Rev, 1988, 88: 563~594.

[10] Schwarz W H E. An introduction to relativistic quantum chemistry. // Barysz M, Ishikawa Y. Relativistic Methods for Chemists. Challenges and Advances in Computational Physics. Vol 10. Dordrecht: Springer Science+Business Media B V , 2010: 1~62.

[11] Hess B A. Relativistic Effects in Heavy-Element Chemistry and Physics. Wiley Series in Theoretical Chemistry. Vol 12. Chichester: John Wiley & Sons, LTD, 2002.

[12] Schwerdtfeger P. Relativistic Electronic Structure Theory. Part 1: Fundamentals. Theoretical and Computational Chemistry. Vol 11. Amsterdam: Elsevier Science B V, 2002.

[13] Schwerdtfeger P. Relativistic Electronic Structure Theory. Part 2: Applications. Theoretical and Computational Chemistry. Vol 14. Amsterdam: Elsevier Science B V, 2004.

[14] Hirao K, Ishikawa Y. Recent Advances in Relativistic Molecular Theory. Recent Advances in Computational Chemistry. Vol 6. London: World Scientific, 2004.

[15] Barysz M, Ishikawa Y. Relativistic Methods for Chemists. Challenges and Advances in Computational Physics. Vol 10. Dordrecht: Springer Science+Business Media B V,

2010.

[16] Pyykkö P. Relativistic Theory of Atoms and Molecules: A Bibliography 1916~1985. Lecture Notes in Chemistry. Vol 41. Berlin: Springer-Verlag, 1986.

[17] Pyykkö P. Relativistic Theory of Atoms and Molecules: II A Bibliography 1986~1992. Lecture Notes in Chemistry. Vol 60. Berlin: Springer-Verlag, 1993.

[18] Pyykkö P. Relativistic Theory of Atoms and Molecules: III A Bibliography 1993~1999. Lecture Notes in Chemistry. Vol 76. Berlin: Springer-Verlag, 2000.

[19] 刘文剑. 相对论量子化学新进展. 化学进展, 2007, 19: 833~851.

[20] 徐光宪，黎乐民，王德民. 量子化学 —— 基本原理和从头计算法 (上中下册). 第二版. 北京: 科学出版社, 2007~2009.

[21] Dyall K G, Faegri K Jr. Introduction to Relativistic Quantum Chemistry. New York: Oxford University Press, 2007.

[22] Reiher M, Wolf A. Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science. Weinheim: Wiley-VCH Verlag GmbH & Co KGaA, 2009.

[23] Grant I P. Relativistic Quantum Theory of Atoms and Molecules: Theory and Computations. Springer Series on Atomic, Optical, and Plasma Physics. New York: Springer Science+Business Media, LLC, 2010.

目　　录

第 1 章　狭义相对论基础

Einstein 于 1905 年创立的狭义相对论是物理学的主要理论基础之一，为核物理和高能物理的发展起到了重要的作用。相对论量子化学也是和狭义相对论密不可分的。为符合狭义相对论的要求，惯性参考系下微观粒子的运动必须重新加以考虑，于是诞生了相对论波动方程。而相对论量子化学就是围绕如何求解相对论波动方程而产生的。

狭义相对论来源于电磁学理论。事实上，可以说正是描述电磁现象的 Maxwell 方程组使人们不得不注意到狭义相对论的存在。自 1890 年以后，Lorentz 和 Poincaré 等作了大量关于电动力学的研究，为狭义相对论的产生奠定了基础。在他们工作的基础上，Einstein 将电动力学领域的研究成果推广到所有物理现象，并强调了光速不变假设的深远影响。他的工作对狭义相对论的产生起到了决定性的作用。

本章主要介绍以下内容：

(1) 狭义相对论的实验基础。

(2) 狭义相对论的两个基本原理 —— 相对性原理和光速不变原理，即 Einstein 提出的两个基本假设，并由此导出时空坐标的 Lorentz 变换式。

(3) 由时空坐标的 Lorentz 变换式，推出惯性参考系之间的速度和质量变换式，并讨论相对论的质量、动量和能量关系。

关于狭义相对论产生的实验基础、历史渊源、大量的推论等详细内容，有兴趣的读者可以参阅 French 编写的 *Special Relativity* 一书[1]。

1.1　19 世纪末的物理学研究

19 世纪末，尽管人们还不能解释所有的实验结果，但是在物理学家中普遍存在一种乐观情绪，认为已有的理论已经足以解释所有的物理现象，不再需要新的理论[2]。然而，也有一些清醒的物理学家已经认识到经典物理学中潜伏的危机。1900 年 4 月 27 日，著名物理学家 Kelvin 勋爵 (本名：W. Thomson) 在英国皇家工学院 (Royal Institute of Great Britain) 的演讲中就指出[3]：经典物理学的上空悬浮着两朵乌云。第一朵乌云涉及当时的光的波动性理论，第二朵乌云则和黑体辐射能量分布公式有关。前者导致了相对论的产生，而后者则催生了量子力学。相对论和量子力学的提出，是 20 世纪物理学的两个重要的里程碑。下面我们就来了解一下狭义

相对论产生的历史背景。

在 1861 年和 1862 年两年时间里，Maxwell 在 *Philosophical Magazine* 相继发表了四篇著名的文章[4−7]，提出了被后人称作 Maxwell 方程组的一组偏微分方程组，描述电场、磁场与电荷密度、电流密度之间的关系。1892 年，Lorentz 提出了 Lorentz 力的概念，并导出 Lorentz 力公式，用以描述电磁场中带电粒子所受到的力。Maxwell 方程组和 Lorentz 力公式构成了经典电动力学的理论基础。从 Maxwell 方程组出发，可以得到光的波动方程，并由此方程得出光在真空中的传播速度 c。按照经典的时空观 (又称 Galileo 时空观)，如果光相对于某一参考系的运动速度为 c，则变换到另一个参考系时，其速度就不可能沿各个方向都为 c。因而 Maxwell 方程组也就只能对一个特殊参考系成立。如果确实是这样，则经典力学中一切惯性参考系等价的相对性原理在电磁现象中就不再成立。另外，当时人们认为既然声波、水波等都是在某个介质中的机械振动的传播现象，则光 (电磁波的一种) 也应该是在某种充满空间的弹性介质 (以太) 内的波动现象。但是，人们无法解释为什么天体可以无摩擦地穿行于“以太”之中，也无法通过实验测出“以太”本身的运动速度。其中一个“寻找以太失败”的著名例子就是用来测量光速沿不同方向差异的 Michelson-Morley 实验，按照 Michelson 和 Morley 的设计，如果能精确测定各个方向光速的差异，就可以确定地球相对于“以太”的运动。然而 Michelson-Morley 实验结果显示光速不依赖于观察者所在的参考系，不存在地球相对于“以太”的运动。自从第一次实验之后，不同的实验工作者还进行过多次的 Michelson-Morley 实验，以不断提高的精确度否定了地球相对于“以太”的运动。

总结新的实验事实后，Einstein 提出了以下两条假设[8]。

1. 相对性原理

所有惯性参考系都是等价的。物理规律和实验结果对于所有惯性参考系都可以表示为相同形式。

2. 光速不变原理

光速和光源的运动无关。真空中的光速相对于任何惯性参考系沿任一方向恒为 c。光速不变性所导致的时空概念是和经典时空观有深刻矛盾的，所有最基本的时空概念，如同时性、距离、时间、速度等都要重新加以探讨。

1.2　间隔不变性

现在我们根据相对论基本原理导出相对论时空坐标变换式。设惯性参考系 Σ' 相对于 Σ 以速度 u 运动，并选 x 轴和 x' 轴沿运动方向，如图 1.1 所示。

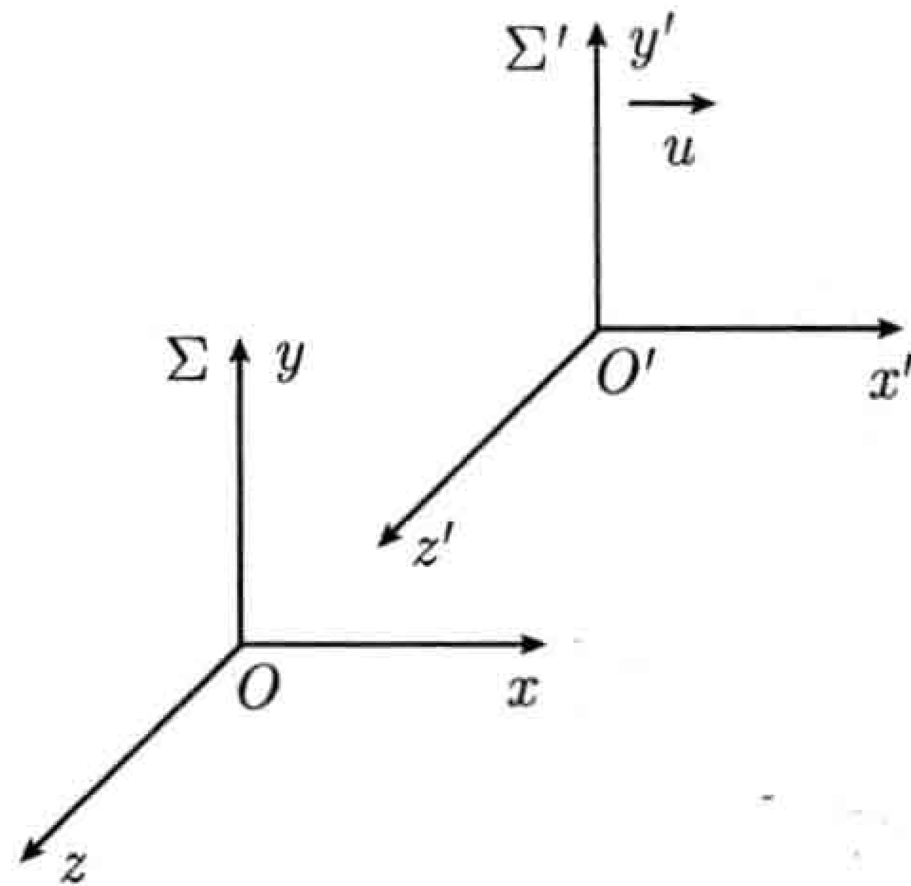


图 1.1 参考系 Σ, Σ' 的相对运动

空间中任意一点 P 在两个参考系 (Σ, Σ') 中的坐标分别用 $(x, y, z), (x', y', z')$ 表示，假定一束光从点 P 发出，由相对论基本原理，光速在两个参考系中均为 c，设光到达原点 O, O' 所需时间分别为 $\Delta t, \Delta t'$，则有

$$\begin{aligned} x^2 + y^2 + z^2 &= c^2(\Delta t)^2 \\ x'^2 + y'^2 + z'^2 &= c^2(\Delta t')^2 \end{aligned} \tag{1.1}$$

比较以上两式可得

$$x^2 + y^2 + z^2 - c^2(\Delta t)^2 = x'^2 + y'^2 + z'^2 - c^2(\Delta t')^2 \tag{1.2}$$

上式即为光速不变性的数学表示，它是相对论时空观的一个基本关系。

定义一个四维矢量 $\boldsymbol{r} = (x, y, z, \mathrm{i}ct)$，其中 x, y, z 为空间坐标，t 为时间坐标。该矢量的长度以矢量的内积 s^2 表示，则有

$$s^2 = \boldsymbol{r} \cdot \boldsymbol{r} = x^2 + y^2 + z^2 - c^2t^2 \tag{1.3}$$

由 (1.2) 式可知，s^2 不因参考系的变换而改变。

实际上，我们可以把光从 P 点发出看成是一个事件，用四维时空坐标 $(x, y, z, 0)$ 表示，光到达原点可看成是另一事件，用 $(0, 0, 0, \mathrm{i}ct)$ 表示，s^2 则表示在某一惯性参考系中所观察到的以上两个事件之间的间隔。所以 s^2 在任何惯性参考系保持恒定的性质又称为间隔不变性。

一般地，两事件 $(x_1, y_1, z_1, \mathrm{i}ct_1)$ 与 $(x_2, y_2, z_2, \mathrm{i}ct_2)$ 的间隔为

$$s^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2(t_1 - t_2)^2 \tag{1.4}$$

在另一参考系上观察这两事件的时空坐标为 $(x_1', y_1', z_1', \mathrm{i}ct_1')$ 和 $(x_2', y_2', z_2', \mathrm{i}ct_2')$，其间隔为

$$s'^2 = (x_1' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 - c^2(t_1' - t_2')^2 \tag{1.5}$$

由间隔不变性有 $s^2 = s'^2$。

间隔是相对论时空观的一个基本概念，由 (1.4) 式，若两事件在同一地点相继发生，间隔就是光速乘以时间的平方的负值。若两事件同时但不在同一地点发生，间隔就是两个事件的空间距离的平方。

1.3 Lorentz 变换

由 Einstein 的相对性原理可知，所有物理公式在不同惯性参考系具有相同的表达形式，这就要求从一个惯性参考系到另一个惯性参考系的时空坐标变换必须是线性变换。下面我们就利用变换的线性和间隔不变性导出相对论时空坐标变换关系。

为简单计算，选两坐标系的 x 轴和 x' 轴都沿 Σ' 相对于 Σ 的运动方向 (见图 1.1)。这种情形下，y 和 z 不变，变换具有特殊形式

$$\begin{aligned} x' &= a_{11}x + a_{12}ct \\ y' &= y \\ z' &= z \\ ct' &= a_{21}x + a_{22}ct \end{aligned} \tag{1.6}$$

将 (1.6) 式代入 (1.2) 式，有

$$\left(a_{11}^2 - a_{21}^2\right)x^2 + 2c\left(a_{11}a_{12} - a_{21}a_{22}\right)xt + \left(a_{12}^2 - a_{22}^2\right)c^2t^2 = x^2 - c^2t^2 \tag{1.7}$$

比较系数得

$$\begin{aligned} a_{11}^2 - a_{21}^2 &= 1 \\ a_{11}a_{12} - a_{21}a_{22} &= 0 \\ a_{12}^2 - a_{22}^2 &= -1 \end{aligned} \tag{1.8}$$

由 (1.8) 式的第一、三式可得

$$a_{11} = \sqrt{1 + a_{21}^2}, \quad a_{22} = \sqrt{1 + a_{12}^2} \tag{1.9}$$

由于参考系 Σ' 和参考系 Σ 的坐标轴和时间轴，即 x 轴和 x' 轴、t 轴和 t' 轴的正方向一致，所以式中 a_{11}，a_{22} 取正值。将 (1.9) 式代入 (1.8) 式的第二个等式，得

$$a_{12}^2 + a_{12}^2a_{21}^2 = a_{21}^2 + a_{21}^2a_{12}^2 \tag{1.10}$$

比较等式两边，得 $a_{12}=a_{21}$，再联合 (1.9) 式，得 $a_{11}=a_{22}$。

考虑 Σ' 的原点 O'，在 Σ 上观察，O' 点以速度 u 沿 x 轴方向运动，因此其坐标为 $x=ut$；O' 点在 Σ' 上的坐标永远是 $x'=0$。由 (1.6) 式中的第一式有

$$0=a_{11}ut+a_{12}ct \tag{1.11}$$

联合 (1.8) 式 ~(1.10) 式，容易求得

$$\begin{aligned} a_{11}=a_{22} &= \frac{1}{\sqrt{1-\dfrac{u^2}{c^2}}} \\ a_{12}=a_{21} &= \frac{-\dfrac{u}{c}}{\sqrt{1-\dfrac{u^2}{c^2}}} \end{aligned} \tag{1.12}$$

将 (1.12) 式代入 (1.6) 式，得相对论时空坐标变换公式

$$\begin{aligned} x' &= \gamma\,(x-ut) \\ y' &= y \\ z' &= z \\ t' &= \gamma\left(t-\frac{u}{c^2}x\right) \end{aligned} \tag{1.13}$$

其中 γ 称为 Lorentz 因子

$$\gamma=\frac{1}{\sqrt{1-\dfrac{u^2}{c^2}}} \tag{1.14}$$

以上坐标变换关系最早由 Lorentz 于 1904 年提出，目的是保持 Maxwell 方程组在不同惯性参考系中的表达形式不变。所以 (1.13) 式又称为 Lorentz 变换式。

在推导上述变换关系时，我们仅考虑了特殊情形，即 Σ' 沿着 Σ 的 x 轴以速度 u 运动。对任意 $\boldsymbol{u}$，则相应的时空坐标变换关系为[9]

$$\begin{aligned} \boldsymbol{r}' &= \gamma\,(\boldsymbol{r}-\boldsymbol{u}t)+(\gamma-1)\,\frac{\boldsymbol{u}\times(\boldsymbol{u}\times\boldsymbol{r})}{u^2} \\ t' &= \gamma\left(t-\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{c^2}\right) \end{aligned} \tag{1.15}$$

式中，$\boldsymbol{r},t$ 和 $\boldsymbol{r}',t'$ 分别表示 Σ,Σ' 两个参考系中的位置和时间变量。

从 Lorentz 变换出发，可以得出一些有意义的推论。

1. 同时的相对性

假定对 Σ 中的观察者而言，两事件 $(x_1,y_1,z_1,\mathrm{i}ct_1)$ 与 $(x_2,y_2,z_2,\mathrm{i}ct_2)$ 同时发生，即 $\Delta t=t_1-t_2=0$，根据 Lorentz 变换 (1.15) 式，对 Σ' 中的观察者，则有

$$\Delta t'=t_1'-t_2'=-\gamma\frac{\boldsymbol{u}\cdot\Delta\boldsymbol{r}}{c^2}\neq 0 \tag{1.16}$$

上式仅当 $\boldsymbol{u}\cdot\Delta\boldsymbol{r}=0$ 时才为零，也就是说若两事件对 Σ 同时则对 Σ' 一般不同时。

2. 运动时钟的延缓

自然界中存在许多物理过程可以作为计时的基准。在 1967 年召开的第 13 届国际度量衡大会上，规定 1 秒的定义是：铯-133 原子的基态两个超精细能级间跃迁辐射的 9 192 631 770 个周期的持续时间。在不同参考系中我们可以用同一物理过程作为计时基准 (又称时钟)，以比较不同参考系中的时间。

例如，分别在参考系 Σ 和 Σ' 观察同一物理过程，测得该过程持续时间分别为 Δt 和 $\Delta t'$，由 Lorentz 变换 (1.15) 式，不难得出 Δt 和 $\Delta t'$ 的关系

$$\Delta t'=\gamma\Delta t \tag{1.17}$$

显然，$\Delta t'>\Delta t$，即运动物体上发生的自然过程比静止物体的同样过程延缓了。物体运动的速度越大，所观察到的它的内部物理过程进行得越缓慢，这就是时间延缓效应。

时间延缓效应已经得到大量实验证明，如携带原子钟的环球飞行实验[10]、高速运动粒子寿命的测定[11]、横向 Doppler 效应实验[12, 13] 等。

3. 运动尺度缩短

现代测量长度也采用自然基准。在 1983 年召开的国际度量衡大会上定义光在真空中行走 1/299 792 458 秒的距离为 1 米。在不同参考系上，都可以用这自然尺度来测量长度。现在我们用 Lorentz 变换式求物体运动时测得的长度与物体静止时长度的关系。

如图 1.2 所示，设物体沿 x 轴方向运动，以固定于物体上的参考系为 Σ'。在 Σ 中同时测量物体两端的坐标，用 x_1, x_2 表示，二者的差值 (x_2-x_1) 定义为 Σ 上测得的物体长度 l。

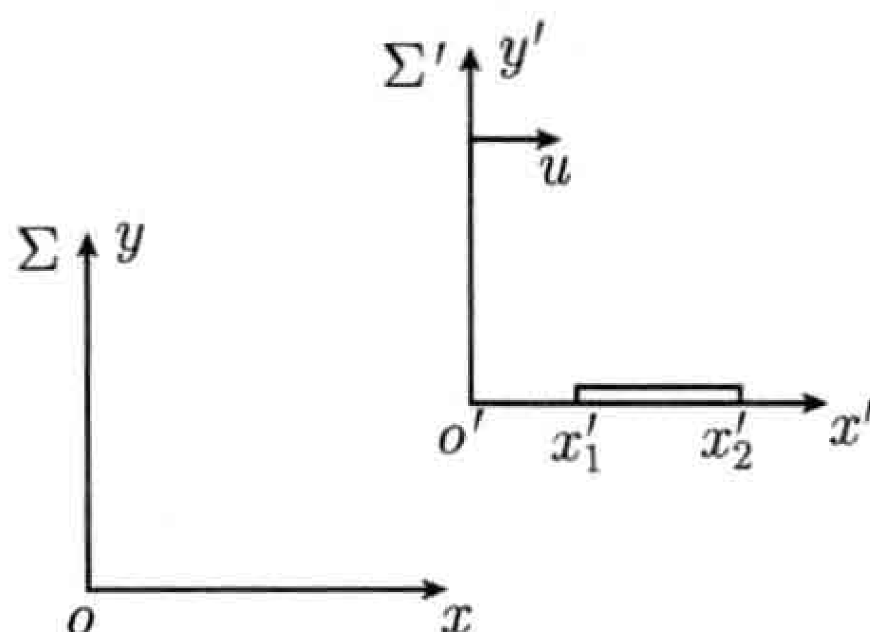


图 1.2　物体在运动参考系的长度

设物体两端在 Σ' 上的坐标为 x_1', x_2'，测得长度为 l_0。由 Lorentz 变换得

$$l_0=x_2'-x_1'=\gamma\left[l+u\left(t_1-t_2\right)\right] \tag{1.18}$$

由于在 Σ 中测量物体两端的坐标时是同时进行的，即 $t_1=t_2$，代入 (1.18) 式得

$$l_0 = l\gamma \tag{1.19}$$

由于 $\gamma > 1$，所以 $l < l_0$，即运动物体长度缩短了。

以上我们证明了在 Σ 上观察固定于 Σ' 上的物体长度缩短了。现在我们在 Σ' 上观察固定于 Σ 上的物体长度，这时要求在 Σ' 上同时测定该物体两端的坐标，即要求 $t_1' = t_2'$。应用 Lorentz 变换得

$$x_2 - x_1 = (x_2' - x_1')\gamma \tag{1.20}$$

现在 $x_2 - x_1$ 为静止长度 l_0，$x_2' - x_1'$ 为运动长度 l，代入 (1.20) 式得

$$l_0 = l\gamma \tag{1.21}$$

与 (1.19) 式一致。也就是说长度缩短效应是相对的。

1.4 速度变换

由 Lorentz 变换可以推出相对论的速度变换公式，设

$$\begin{aligned} v_x &= \frac{\mathrm{d}x}{\mathrm{d}t} \\ v_y &= \frac{\mathrm{d}y}{\mathrm{d}t} \\ v_z &= \frac{\mathrm{d}z}{\mathrm{d}t} \end{aligned} \tag{1.22}$$

为物体相对于 Σ 的速度。考虑 Σ' 沿 Σ 的 x 轴以速度 u 运动。对 Lorentz 变换 (1.13) 式微分，有

$$\begin{aligned} \mathrm{d}x' &= \gamma(\mathrm{d}x - u\mathrm{d}t) \\ \mathrm{d}t' &= \gamma\left(1 - \frac{u}{c^2}v_x\right)\mathrm{d}t \end{aligned} \tag{1.23}$$

两式相除得

$$v_x' = \frac{v_x - u}{1 - \dfrac{uv_x}{c^2}} \tag{1.24}$$

同样可以求得

$$\begin{aligned} v_y' &= \frac{v_y}{\gamma\left(1 - \dfrac{uv_x}{c^2}\right)} \\ v_z' &= \frac{v_z}{\gamma\left(1 - \dfrac{uv_x}{c^2}\right)} \end{aligned} \tag{1.25}$$

(1.24) 式和 (1.25) 式即为不同惯性参考系之间的相对论速度变换。

在非相对论极限条件下，此时 $v \ll c$, $u \ll c$，(1.24) 式和 (1.25) 式变为

$$\begin{aligned} v'_x &\approx v_x - u \\ v'_y &\approx v_y \\ v'_z &\approx v_z \end{aligned} \tag{1.26}$$

此时相对论速度变换过渡到经典速度变换 —— Galileo 变换。

以上我们讨论的是参考系 Σ' 沿参考系 Σ 的 x 轴运动的特殊情形，即 $u_y = u_z = 0$。考虑一般情况，对任一 $\boldsymbol{u}$，由 (1.15) 式，采取与上类似的步骤，得相对论速度变换关系

$$\boldsymbol{v}' = \frac{\boldsymbol{v} - \gamma \boldsymbol{u} + \dfrac{\gamma - 1}{u^2}\boldsymbol{r}\,(\boldsymbol{u} \cdot \boldsymbol{v})}{\gamma\left(1 - \dfrac{\boldsymbol{u} \cdot \boldsymbol{v}}{c^2}\right)} \tag{1.27}$$

相对论速度的上限　真空中的光速 c 是目前已知的速度极限，从相对论的速度变换 (1.24) 式和 (1.25) 式出发，容易证明，当 $v_x < c$，且 $u < c$ 时，v'_x 也永远小于 c。即如果物体在某一惯性参考系的速度小于 c，则在所有其他惯性参考系中测得的该物体的速度都将小于 c。也就是说没有任何物体或信息运动的速度可以超过真空中的光速 c。对比经典速度变换公式 (1.26) 式我们发现，在经典物理中速度是没有上限的。

1.5　质量变换

让我们假想一个简单的完全弹性碰撞实验，如图 1.3 所示。Σ' 的观察者对实验的记录为，质量为 m' 的一个粒子沿 y' 轴运动，速度为 v'_y，粒子与一堵无限长的墙 ($x'z'$ 平面) 碰撞后原路返回，沿 $-y'$ 轴运动，速度仍然是 v'_y。碰撞前后粒子的动量变化为

$$\Delta p' = 2m'v'_y \tag{1.28}$$

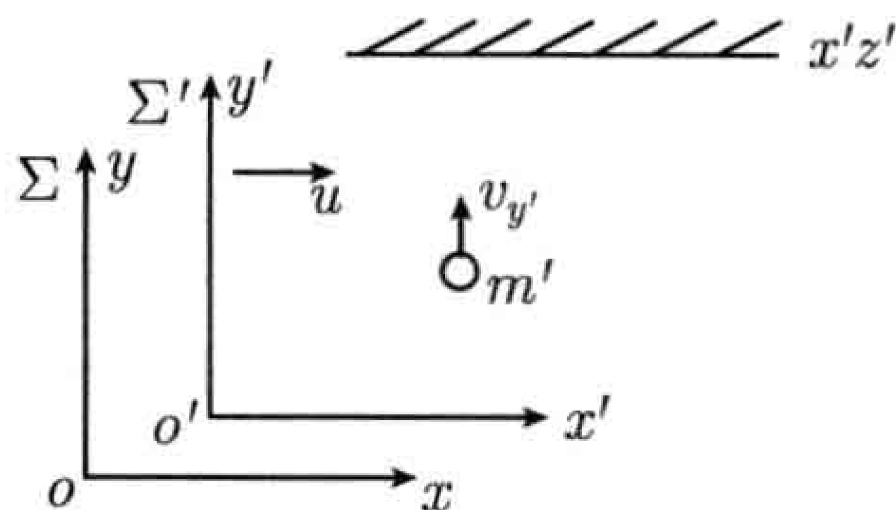


图 1.3　粒子 m' 和无限长墙 $(x'z')$ 的完全弹性碰撞

对 Σ 的观察者，粒子的动量变化为

$$\Delta p = 2mv_y \tag{1.29}$$

完全弹性碰撞遵守动量守恒定律，所以碰撞前后粒子动量的变化等于墙的动量的变化 (Δp_{w}) 的负值，和参考系的选择无关，即

$$\begin{aligned}\Delta p &= -\Delta p_{\mathrm{w}} \\ \Delta p' &= -\Delta p'_{\mathrm{w}}\end{aligned} \tag{1.30}$$

式中，$\Delta p_{\mathrm{w}}, \Delta p'_{\mathrm{w}}$ 分别代表 Σ, Σ' 中测得的墙的动量变化。我们已经假定墙是无限长的，于是有 $\Delta p_{\mathrm{w}} = \Delta p'_{\mathrm{w}}$，由 (1.28) 式 ~(1.30) 式得

$$m'v'_y = mv_y \tag{1.31}$$

联合 (1.24) 式和 (1.25) 式，考虑到 $v'_x = 0$，有

$$v'_y = \gamma v_y \tag{1.32}$$

将该结果代入 (1.31) 式，得相对论质量变换关系

$$m = m'\gamma \tag{1.33}$$

式中，m' 又称为静止质量，习惯上用 m_0 表示。(1.33) 式表明，物体的质量随着物体运动速度的增大而增大，当物体运动的速度接近光速时，物体的质量也趋于无穷。

1.6 相对论能量

根据 Newton 第二定律，作用到物体上的力和物体的加速度之间的关系为

$$\boldsymbol{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{u}) \tag{1.34}$$

物体若在 $\boldsymbol{F}$ 的作用下发生一无限小的位移 $\mathrm{d}x$，$\boldsymbol{F}$ 所做的功可写为

$$F\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}(mu)\,\mathrm{d}x \tag{1.35}$$

假定物体的初速度为零，在 $\boldsymbol{F}$ 的作用下开始运动，末速度为 $\boldsymbol{u}_{\mathrm{f}}$，则 $\boldsymbol{F}$ 所做的总功可写为以下积分形式

$$\begin{aligned}T &= \int_0^{u_{\mathrm{f}}} F\mathrm{d}x \\ &= \int_0^{u_{\mathrm{f}}} \frac{\mathrm{d}}{\mathrm{d}t}(mu)\,\mathrm{d}x \\ &= \int_0^{u_{\mathrm{f}}} u\mathrm{d}(mu) \\ &= \int_0^{u_{\mathrm{f}}} \left(u^2\mathrm{d}m + mu\mathrm{d}u\right)\end{aligned} \tag{1.36}$$

将 (1.14) 式代入 (1.33) 式后方程两边平方，得

$$u^2 = c^2\left(1 - \frac{m_0^2}{m^2}\right) \tag{1.37}$$

对上式微分，有

$$u\mathrm{d}u = \frac{c^2 m_0^2}{m^3}\mathrm{d}m \tag{1.38}$$

将 (1.37) 式和 (1.38) 式代入 (1.36) 式，整理得

$$\begin{aligned} T &= \int_0^{u_\mathrm{f}} c^2 \mathrm{d}m \\ &= m_\mathrm{f}c^2 - m_0 c^2 \end{aligned} \tag{1.39}$$

在能量守恒的条件下，$\boldsymbol{F}$ 所做的功必然转化为体系能量的增加，对刚性物体来说，即为物体动能 (T) 的增加。如果我们将 m_0c^2 解释为静止能量，用 E_0 表示

$$E_0 = m_0 c^2 \tag{1.40}$$

则物体的总能量 (E) 可写为

$$E = mc^2 = T + m_0 c^2 \tag{1.41}$$

将 (1.37) 式代入 $p^2 = m^2u^2$ 或 $\boldsymbol{u}\cdot\boldsymbol{p} = mu^2$，再运用 (1.41) 式可得相对论能量、动量和质量的两条重要关系式，即

$$E^2 = p^2c^2 + m_0^2c^4 \tag{1.42}$$

和

$$E = \boldsymbol{u}\cdot\boldsymbol{p} + \frac{m_0c^2}{\gamma} \tag{1.43}$$

(1.40) 式表示物体的静止质量和静止能量的关系，称为质能关系式。静止能量的揭示是相对论的重要推论之一。它告诉我们一定质量的物体具有一定的内部运动能量。反过来，具有一定内部运动能量的物体就表现出一定的惯性质量。质能关系式在原子核和粒子物理中被大量实验很好地证实，它是原子能利用的主要理论根据。

1.7　相对论动量

将 (1.42) 式重新整理，得

$$p^2 - \frac{E^2}{c^2} = -m_0^2c^2 \tag{1.44}$$

以上等式对任何参考系都成立。由此，我们可以定义动量的四维矢量形式

$$\boldsymbol{p} = \left(p_x, p_y, p_z, \frac{\mathrm{i}E}{c}\right) = (p_x, p_y, p_z, \mathrm{i}mc) \tag{1.45}$$

四维矢量 $\boldsymbol{p}$ 的长度定义为内积

$$\boldsymbol{p} \cdot \boldsymbol{p} = p^2 - \frac{E^2}{c^2} = -m_0c^2 \tag{1.46}$$

由上式容易看出，和 1.2 节定义的四维矢量 $\boldsymbol{r}$ 一样，$\boldsymbol{p}$ 的长度也具有 Lorentz 变换不变性，即不随参考系的变换而改变。

参 考 文 献

[1] French A P. Special Relativity. MIT Introductory Physics Series. Cambridge: Massachussets Institute of Technology, 1968.

[2] Wróblewski A K. Physics in 1900. Acta Phys Pol B, 2000, 31: 179∼196.

[3] Kelvin W T. I. Nineteenth century clouds over the dynamical theory of heat and light. Phil Mag S 6, 1901, 2: 1∼41.

[4] Maxwell J C. XXV. On physical lines of force. Phil Mag S 4, 1861, 21: 161∼175.

[5] Maxwell J C. XXV. On physical lines of force. Phil Mag S 4, 1861, 21: 281∼291.

[6] Maxwell J C. III. On physical lines of force. Phil Mag S 4, 1862, 23: 12∼24.

[7] Maxwell J C. III. On physical lines of force. Phil Mag S 4, 1862, 23: 85∼95.

[8] Einstein A. Zur Elektrodynamik bewegter Körper. Ann Physik, 1905, 17: 891∼921.

[9] Jackson J D. Classical Electrodynamics. New York: John Wiley & Sons Inc, 1975: 517.

[10] Hafele J C, Keating R E. Around-the-world atomic clocks: predicted relativistic time gains. Science, 1972, 177: 166∼168.

[11] Greenberg A J, Ayres D S, Cormack A M, et al. Charged-pion lifetime and a limit on a fundamental length. Phys Rev Lett, 1969, 23: 1267∼1270.

[12] Ives H E, Stilwell G R. An experimental study of the rate of a moving atomic clock. II. J Opt Soc Am, 1941, 31: 369∼374.

[13] Ives H E, Stilwell G R. An experimental study of the rate of a moving atomic clock. J Opt Soc Am, 1938, 28: 215∼219.

第 2 章　相对论波动方程

一个微观粒子的量子态可以用波函数 $\Psi(\boldsymbol{r},t)$ 来描述，当 $\Psi(\boldsymbol{r},t)$ 确定后，粒子的任何一个可观测量的平均值以及它取各种可能值的几率都完全确定。那么如何解决波函数随时间演化以及如何求解波函数呢？

1926 年，Schrödinger 发表了描述波函数变化的非相对论波动方程，即 Schrödinger 方程，解决了非相对论条件下波函数随时间演化的问题；同年，Klein 和 Gordon 提出了第一个相对论波动方程 —— Klein-Gordon 方程，但该方程仅适用于自旋为零的粒子；两年后 Dirac 发表了描述自旋为 1/2 的粒子的相对论波动方程 —— Dirac 方程。

本章将首先介绍自由粒子的 Klein-Gordon 方程，包括方程的导出、方程的非相对论极限、由方程定义的几率密度和几率流密度等。由于自旋为零的粒子不是化学工作者的主要研究对象，我们将不涉及多粒子的 Klein-Gordon 方程，也不考虑外加电磁场的作用，对此有兴趣的读者可以阅读 Greiner 撰写的 *Relativistic Quantum Mechanics: Wave Equations*[1]。本章介绍的重点是 Dirac 方程，在给出 Dirac 方程及有关概念后，我们将主要讨论自由电子和氢原子的 Dirac 方程严格解。

2.1　自由粒子的 Klein-Gordon 方程

在非相对论条件下，自由粒子的 Schrödinger 方程为

$$\mathrm{i}\hbar\frac{\partial\Psi(\boldsymbol{r},t)}{\partial t}=\hat{H}\,\Psi(\boldsymbol{r},t) \tag{2.1}$$

其中自由粒子的 Hamilton 算符 ($\hat{H}$) 形式与经典分析力学中的 Hamilton 算符一致，即

$$\hat{H}=\frac{\hat{\boldsymbol{p}}^2}{2m} \tag{2.2}$$

式中，$\dfrac{\hat{\boldsymbol{p}}^2}{2m}$ 对应经典自由粒子的能量 (E)。第一个相对论的运动方程正是仿照这种对应方式，利用 (1.42) 式将能量改为相对论关系而得到的，即

$$\mathrm{i}\hbar\frac{\partial\Psi(\boldsymbol{r},t)}{\partial t}=\left[c\left(m_0^2c^2+\hat{\boldsymbol{p}}^2\right)^{1/2}\right]\Psi(\boldsymbol{r},t) \tag{2.3}$$

或

$$\left(\mathrm{i}\hbar\frac{\partial}{\partial t}\right)^2\Psi(\boldsymbol{r},t)=\left(m_0^2c^4+\hat{\boldsymbol{p}}^2c^2\right)\Psi(\boldsymbol{r},t) \tag{2.4}$$

这个方程分别由 Klein 和 Gordon 独立提出，被后人称为自由粒子的 Klein-Gordon 方程。

方程 (2.4) 式是一个经典的波动方程，其标准解为

$$\Psi(\boldsymbol{r},t)=\exp\left[\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)\right] \tag{2.5}$$

将 (2.5) 式代入 (2.4) 式，可求得能量 E

$$E=\pm E_p=\pm\sqrt{m_0^2c^4+p^2c^2} \tag{2.6}$$

由上可以看出，能量有负值解，而且没有下限。

2.1.1 Klein-Gordon 方程的非相对论极限

为求得 Klein-Gordon 方程的非相对论极限，我们将能量写为两项的和，即 $E=E'+m_0c^2$，于是满足 Klein-Gordon 方程的波函数 (2.5) 式可重写为

$$\begin{aligned}\Psi(\boldsymbol{r},t)&=\exp\left[\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)\right]\\&=\Phi(\boldsymbol{r},t)\exp\left(-\frac{\mathrm{i}}{\hbar}m_0c^2t\right)\end{aligned} \tag{2.7}$$

式中，Φ 不包含静止能量，且 Φ 满足

$$\mathrm{i}\hbar\frac{\partial\Phi}{\partial t}=E'\Phi \tag{2.8}$$

考虑非相对论极限条件，即粒子速度远小于光速 ($v\ll c$)，我们将相对论能量表达式 (见 (1.42) 式) 作 Taylor 展开

$$E=m_0c^2\left(1+\frac{v^2}{2c^2}-\frac{v^4}{8c^4}+\cdots\right) \tag{2.9}$$

由上式可知，当 $v\ll c$ 时，有

$$E'=E-m_0c^2\ll m_0c^2 \tag{2.10}$$

代入 (2.8) 式，有

$$\left|\mathrm{i}\hbar\frac{\partial\Phi}{\partial t}\right|=|E'\Phi|\ll|m_0c^2\Phi| \tag{2.11}$$

将波函数 Ψ 对时间求偏微分，并考虑到 (2.11) 式，得

$$\begin{aligned}\frac{\partial\Psi}{\partial t}&=\left(\frac{\partial\Phi}{\partial t}-\mathrm{i}\frac{m_0c^2}{\hbar}\Phi\right)\exp\left(-\frac{\mathrm{i}}{\hbar}m_0c^2t\right)\\&\approx-\mathrm{i}\frac{m_0c^2}{\hbar}\Phi\exp\left(-\frac{\mathrm{i}}{\hbar}m_0c^2t\right)\end{aligned} \tag{2.12}$$

将波函数 Ψ 对时间求二阶偏微分，并运用上式的结果，有

$$\begin{aligned}\frac{\partial^2 \Psi}{\partial t^2} &= \frac{\partial}{\partial t}\left[\left(\frac{\partial \Phi}{\partial t} - \mathrm{i}\frac{m_0c^2}{\hbar}\Phi\right)\exp\left(-\frac{\mathrm{i}}{\hbar}m_0c^2t\right)\right] \\ &\approx \left(-2\mathrm{i}\frac{m_0c^2}{\hbar}\frac{\partial \Phi}{\partial t} - \frac{m_0^2c^4}{\hbar^2}\Phi\right)\exp\left(-\frac{\mathrm{i}}{\hbar}m_0c^2t\right)\end{aligned} \tag{2.13}$$

将 (2.13) 式代回到 Klein-Gordon 方程 (见 (2.4) 式)，整理得

$$\mathrm{i}\hbar\frac{\partial \Phi}{\partial t} = -\frac{\hbar^2}{2m_0}\nabla^2\Phi \tag{2.14}$$

上式即为自旋为零的粒子的 Schrödinger 方程。由此我们推断 Klein-Gordon 方程描述的是自旋为零的粒子的相对论波动方程。

2.1.2 几率密度和几率流密度

将 $\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla$ 代入 (2.4) 式，整理得

$$\left(\frac{\partial^2}{c^2\partial t^2} - \nabla^2 + \frac{m_0^2c^2}{\hbar^2}\right)\Psi(\boldsymbol{r},t) = 0 \tag{2.15}$$

算符 ∇ 的定义见附录 B。(2.15) 式的复共轭形式为

$$\left(\frac{\partial^2}{c^2\partial t^2} - \nabla^2 + \frac{m_0^2c^2}{\hbar^2}\right)\Psi^*(\boldsymbol{r},t) = 0 \tag{2.16}$$

将 (2.15) 式和 (2.16) 式分别左乘 Ψ^* 和 Ψ 后再相减，得

$$\frac{\partial}{\partial t}\left[\frac{\mathrm{i}\hbar}{2m_0c^2}\left(\Psi^*\frac{\partial \Psi}{\partial t} - \Psi\frac{\partial \Psi^*}{\partial t}\right)\right] - \frac{\mathrm{i}\hbar}{2m_0}\left[\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*\right] = 0 \tag{2.17}$$

进一步将式中第二项变形，有

$$\frac{\partial}{\partial t}\left[\frac{\mathrm{i}\hbar}{2m_0c^2}\left(\Psi^*\frac{\partial \Psi}{\partial t} - \Psi\frac{\partial \Psi^*}{\partial t}\right)\right] - \frac{\mathrm{i}\hbar}{2m_0}\nabla\cdot\left[\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\right] = 0 \tag{2.18}$$

将 (2.18) 式和连续性方程

$$\nabla\cdot\hat{\boldsymbol{j}} + \frac{\partial \rho}{\partial t} = 0 \tag{2.19}$$

对比，不难得到几率流密度 ($\hat{\boldsymbol{j}}$) 和几率密度 (ρ) 的表达式

$$\rho(\boldsymbol{r},t) = \frac{\mathrm{i}\hbar}{2m_0c^2}\left(\Psi^*\frac{\partial \Psi}{\partial t} - \Psi\frac{\partial \Psi^*}{\partial t}\right) \tag{2.20}$$

$$\hat{\boldsymbol{j}}(\boldsymbol{r},t) = -\frac{\mathrm{i}\hbar}{2m_0}\left(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\right) \tag{2.21}$$

上式两边分别乘以单位电荷 e，即得我们熟知的电荷密度和电流密度。几率流密度的表达式，即 (2.21) 式，和非相对论的表达式类似，容易证明 $\hat{j}$ 的值始终是实数。我们再来看几率密度 ρ，在某一时刻 t，$\varPsi$ 和 $\dfrac{\partial \varPsi}{\partial t}$ 可以并不相关，则由 (2.20) 式知几率密度 ρ 的值可能大于零，也可能小于零。而在非相对论条件下，几率密度表达式为

$$\rho(\boldsymbol{r}, t) = \varPsi\varPsi^* \geqslant 0 \tag{2.22}$$

为什么从 Klein-Gordon 方程得到的几率密度会出现小于零的情况呢？为回答这个问题，我们将 (2.4) 式中的 $\varPsi$ 重写为

$$\varPsi(\boldsymbol{r}, t) = \varphi + \chi \tag{2.23}$$

定义一个新函数 $\varPhi$

$$\varPhi(\boldsymbol{r}, t) = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \tag{2.24}$$

读者可以自己证明，$\varPhi(\boldsymbol{r}, t)$ 满足 Schrödinger 方程

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\varPhi(\boldsymbol{r}, t) = \hat{H}\varPhi(\boldsymbol{r}, t) \tag{2.25}$$

其中

$$\hat{H} = (\hat{\sigma}_z + \mathrm{i}\hat{\sigma}_y)\frac{\hat{\boldsymbol{p}}^2}{2m_0} + \hat{\sigma}_z m_0 c^2 \tag{2.26}$$

$\hat{\boldsymbol{\sigma}}$ 是一个矢量算符，表示为 Pauli 创立的一组 2×2 复数矩阵，称为 Pauli 算符。

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.27}$$

将 (2.23) 式代入几率密度表达式 (2.20) 式并利用 (2.25) 式，得电荷密度为

$$e\rho(\boldsymbol{r}, t) = e(\varphi^*\varphi - \chi^*\chi) = e\varPhi^\dagger\hat{\sigma}_z\varPhi \tag{2.28}$$

将上式对整个空间积分，得

$$\begin{aligned} \iiint_{-\infty}^{+\infty} e\rho(\boldsymbol{r}, t)\,dx\,dy\,dz &= \iiint_{-\infty}^{+\infty} e\varPhi^\dagger\hat{\sigma}_z\varPhi\,dx\,dy\,dz \\ &= \pm e \end{aligned} \tag{2.29}$$

由此我们发现，(2.20) 式出现负值是由粒子所带电荷 $(-e)$ 引起的。综合 2.1.1 节的结果，我们得知 Klein-Gordon 方程 (2.4) 式描述的是自旋为零的粒子的运动，该粒子所带电荷为 $\pm e$。

2.2 Dirac 方程

2.2.1 自由粒子的 Dirac 方程

在 Klein-Gordon 方程发表后，由于当时还没有发现自旋为零的介子 π^+ 和 π^-，所以人们始终怀疑其正确性。于是 Dirac 开始寻找相对论运动方程的另一种形式，他希望它和非相对论的 Schrödinger 方程一样，都是一个对时间的一阶导数方程。Dirac 假设自由粒子的相对论运动方程应取下列形式

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t)=\hat{h}_{\mathrm{D}}\Psi(\boldsymbol{r},t) \tag{2.30}$$

此式被后人称为自由粒子的 Dirac 方程，$\hat{h}_{\mathrm{D}}$ 是自由粒子的 Dirac Hamilton 算符，Dirac 将其写为

$$\hat{h}_{\mathrm{D}}=c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}+\hat{\beta}m_0c^2 \tag{2.31}$$

式中，$\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 是与时间和位置无关的待定常量。显然 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 不可能是简单常数，否则一个简单的空间转动就会改变 Dirac 方程的形式 ——这就违背了 Einstein 的相对性原理 (见第 1 章)。我们假定 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 是矩阵，并赋予算符标志 $\wedge$，为保证 Dirac Hamilton 算符的厄米性，$\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 也必须是厄米算符，即矩阵的本征值为实数。波函数 $\Psi(\boldsymbol{r},t)$ 此时不再是一个标量，而是一个列向量

$$\Psi=\begin{pmatrix}\psi_1(\boldsymbol{r},t)\\ \psi_2(\boldsymbol{r},t)\\ \cdots\\ \psi_N(\boldsymbol{r},t)\end{pmatrix} \tag{2.32}$$

为使 Dirac 方程两边列向量的维数一致，$\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 显然应该是 $N\times N$ 的方阵。

此外，自由粒子 Dirac 方程的能量解还必须满足相对论能量表达式，即 $E^2=p^2c^2+m_0^2c^4$ 。由 2.1 节我们已经知道 Klein-Gordon 方程的能量解符合上述要求，于是 Dirac 要求 (2.30) 式也必须同时满足 Klein-Gordon 方程。将 (2.30) 式等号右边的项移到左边，再左乘 $\mathrm{i}\hbar\frac{\partial}{\partial t}+\hat{h}_{\mathrm{D}}$，得

$$\left(-\hbar^2\frac{\partial^2}{\partial t^2}-\hat{h}_{\mathrm{D}}^2\right)\Psi(\boldsymbol{r},t)=0 \tag{2.33}$$

为了与自由粒子的 Klein-Gordon 方程 (2.4) 式一致，则待定常量应满足

$$\begin{aligned}\alpha_x^2=\alpha_y^2=\alpha_z^2=\beta^2&=1\\ \alpha_i\alpha_j+\alpha_j\alpha_i&=0\qquad(i\neq j)\\ \alpha_i\beta+\beta\alpha_i&=0\end{aligned} \tag{2.34}$$

由上式中的第一式，再考虑到 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 为厄米算符，可得 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 的本征值只能是 ±1 。于是我们可以将 $\hat{\boldsymbol{\alpha}}$ 写成如下形式

$$\begin{pmatrix} \pm1 & 0 & 0 & \cdots & 0 \\ 0 & \pm1 & 0 & \cdots & 0 \\ 0 & 0 & \pm1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \pm1 \end{pmatrix} \tag{2.35}$$

由 (2.34) 式中第二、三式知 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 是反对易的，得 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 的迹，即矩阵对角元之和必为零，联合 (2.35) 式可知矩阵的维数 N 必为偶数。

对最小的偶数 2，最多只能构造三个反对易的 2×2 矩阵，即 Pauli 矩阵 (2.27) 式，而我们需要四个 2×2 矩阵，即 $\alpha_x,\alpha_y,\alpha_z,\beta$。于是得到满足 (2.34) 式的最小矩阵维数为 4。Dirac 将 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 写成如下形式

$$\hat{\boldsymbol{\alpha}} = \begin{pmatrix} 0 & \hat{\boldsymbol{\sigma}} \\ \hat{\boldsymbol{\sigma}} & 0 \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} \boldsymbol{I}_2 & 0 \\ 0 & \boldsymbol{I}_2 \end{pmatrix} \tag{2.36}$$

式中，$\hat{\boldsymbol{\sigma}}$ 为 Pauli 矩阵；$\boldsymbol{I}_2$ 为 2×2 单位矩阵。将 (2.27) 式代入 (2.36) 式，得

$$\begin{gathered}\hat{\alpha}_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \hat{\alpha}_y = \begin{pmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \end{pmatrix} \\ \hat{\alpha}_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\end{gathered} \tag{2.37}$$

由此我们知道波函数 $\Psi(\boldsymbol{r},t)$ 是一个包含四个分量的矢量，所以相对论波函数也被称为四分量波函数。需要指出的是，以上 Dirac 关于 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 的选择不是唯一的，容易证明，对 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 的任意一个酉变换 $(\hat{U})$，即 $\hat{\alpha}'_i=\hat{U}\hat{\alpha}_i\hat{U}^{-1}$，$\hat{\beta}'=\hat{U}\hat{\beta}\hat{U}^{-1}$，新矩阵 $\hat{\alpha}'_i$，$\hat{\beta}'$ 也满足上述所有要求。

2.2.2 几率密度和几率流密度

在 2.1.2 节中我们介绍了如何从 Klein-Gordon 方程得到连续性方程，即 (2.19) 式，从而推出几率密度和几率流密度的表达式。从 Dirac 方程出发，采取类似步骤

可以得到如下连续性方程

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\left(\Psi^{\dagger}\Psi\right)+\mathrm{i}c\hbar\nabla\cdot\left(\Psi^{\dagger}\hat{\boldsymbol{\alpha}}\Psi\right)=0 \tag{2.38}$$

由上可得几率密度和几率流密度表达式为

$$\begin{aligned}\rho&=\Psi^{\dagger}\Psi=\sum_{i=1}^{4}\psi_i^*(\boldsymbol{r},t)\psi_i(\boldsymbol{r},t)\geqslant 0\\ \hat{\boldsymbol{j}}&=c\Psi^{\dagger}\hat{\boldsymbol{\alpha}}\Psi\end{aligned} \tag{2.39}$$

与由 Klein-Gordon 方程得到的几率密度有可能小于零的情况不同，从 Dirac 方程得到的几率密度总是非负数，与非相对论 Schrödinger 方程的结论一致。

2.2.3　电磁场中单粒子的 Dirac 方程

由经典电动力学可知，处于电磁场 $(\boldsymbol{a},\phi)$ 中粒子的经典 Hamilton 量为

$$\hat{H}=\frac{\hat{\boldsymbol{\pi}}^2}{2m_0}+q\phi \tag{2.40}$$

$\hat{\boldsymbol{\pi}}$ 为动能动量，写为

$$\hat{\boldsymbol{\pi}}=\hat{\boldsymbol{p}}-\frac{q}{c}\boldsymbol{a} \tag{2.41}$$

式中，$\hat{\boldsymbol{p}},\boldsymbol{a},\phi$ 分别表示正则动量、电磁场的矢量势和标量势；q 为粒子所带电荷。我们用 $\hat{\boldsymbol{\pi}}$ 代替自由粒子 Dirac 方程中的 $\hat{\boldsymbol{p}}$，再将粒子在电磁场中的能量 $q\phi$ 加入，即得电磁场中单粒子的 Dirac 方程

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t)=\left(c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{\pi}}+\hat{\beta}m_0c^2+q\phi\right)\Psi(\boldsymbol{r},t) \tag{2.42}$$

等式右边括号内的各项之和即为单粒子的 Dirac Hamilton 量。

现假定该单粒子为电子，而外电磁场 $(\boldsymbol{a},\phi)$ 由原子核 (原子序数为 Z_λ) 产生，并忽略核的运动，则有 $\boldsymbol{a}_\lambda=0,\phi_\lambda=\dfrac{Z_\lambda e}{4\pi\epsilon_0 r}$，代入 (2.42) 式，得单电子的 Dirac Hamilton 算符为

$$\hat{h}_{\mathrm{D}}=c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}+\hat{\beta}m_0c^2-\sum_{\lambda}\frac{Z_\lambda e^2}{4\pi\epsilon_0 r} \tag{2.43}$$

式中，ϵ_0 为真空介电常量。

对多电子体系 (如原子或分子)，在核冻结近似及不存在外加电磁场的情况下，体系的 Dirac Hamilton 算符为

$$\hat{H}=\sum_{i}^{n}\hat{h}_{\mathrm{D}}(i)+\sum_{i<j}^{n}\hat{g}(i,j)+\sum_{\lambda<\mu}^{N}\frac{Z_\lambda Z_\mu}{4\pi\epsilon_0 r_{\lambda\mu}} \tag{2.44}$$

式中，n 代表体系电子数目；N 代表原子核数目。式中第二项表示电子–电子相互作用，具体表达形式我们将在第 3 章讨论。第三项为核–核 Coulomb 相互作用。

2.2.4 Dirac 方程的非相对论极限

为方便得到 Dirac 方程的非相对论极限，我们将四分量的波函数写成两个二分量形式

$$\Psi = \begin{pmatrix} \tilde{\varphi_1} \\ \tilde{\varphi_2} \end{pmatrix} \tag{2.45}$$

进一步，与 2.1.1 节讨论 Klein-Gordon 方程非相对论极限时的方法类似，将静止能量 (m_0c^2) 分离出来，令

$$\begin{pmatrix} \tilde{\varphi_1} \\ \tilde{\varphi_2} \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \exp\left(-\frac{\mathrm{i}}{\hbar} m_0 c^2 t\right) \tag{2.46}$$

将上式代入 Dirac 方程 (2.42) 式，得

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \begin{pmatrix} c\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}}\varphi_2 \\ c\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}}\varphi_1 \end{pmatrix} + q\phi\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} - 2m_0c^2\begin{pmatrix} 0 \\ \varphi_2 \end{pmatrix} \tag{2.47}$$

以上方程实际上包含两个方程，且方程对应的能量解 E' 不包括静止能量，即 $E' = E - m_0c^2$。考虑非相对论极限 $(v \ll c)$，此时 $E' = E - m_0c^2 \ll m_0c^2$(见 (2.10) 式)，也就是说我们可以将粒子的动能和势能均远小于 m_0c^2 作为非相对论极限条件，于是有以下近似成立

$$\begin{aligned} \left|\mathrm{i}\hbar\frac{\partial\varphi_2}{\partial t}\right| &\ll \left|m_0c^2\varphi_2\right| \\ \left|q\phi\varphi_2\right| &\ll \left|m_0c^2\varphi_2\right| \end{aligned} \tag{2.48}$$

将以上近似用于 (2.47) 式的第二个方程，得

$$\varphi_2 = \frac{\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}}}{2m_0c}\varphi_1 \tag{2.49}$$

将 φ_2 的表达式代入 (2.47) 式的第一个方程，得到关于 φ_1 的方程

$$\mathrm{i}\hbar\frac{\partial\varphi_1}{\partial t} = \frac{(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}})(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}})}{2m_0}\varphi_1 + q\phi\varphi_1 \tag{2.50}$$

由附录 B 中的 (B.3) 式，知 Pauli 矢量 $\boldsymbol{\sigma}$ 与任意矢量 $\boldsymbol{B}, \boldsymbol{C}$ 存在以下关系

$$(\boldsymbol{\sigma} \cdot \boldsymbol{C})(\boldsymbol{\sigma} \cdot \boldsymbol{B}) = \boldsymbol{I}_2\boldsymbol{C} \cdot \boldsymbol{B} + \mathrm{i}\boldsymbol{\sigma} \cdot (\boldsymbol{C} \times \boldsymbol{B}) \tag{2.51}$$

式中，$\boldsymbol{I}_2$ 为 2×2 单位矩阵。上式也称 Pauli 矢量规则。用 $\hat{\boldsymbol{\pi}}$ 代替式中 $\boldsymbol{B}, \boldsymbol{C}$ 得

$$(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}})(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\pi}}) = \left(\hat{\boldsymbol{p}} - \frac{q}{c}\boldsymbol{a}\right)^2 - \frac{q\hbar}{c}\hat{\boldsymbol{\sigma}} \cdot (\nabla \times \boldsymbol{a}) \tag{2.52}$$

将上式代入 (2.50) 式，有

$$\mathrm{i}\hbar\frac{\partial\varphi_1}{\partial t}=\left[\frac{\left(\hat{\boldsymbol{p}}-\dfrac{q}{c}\boldsymbol{a}\right)^2}{2m_0}-\frac{q\hbar}{2m_0c}\hat{\boldsymbol{\sigma}}\cdot(\nabla\times\boldsymbol{a})+q\phi\right]\varphi_1 \tag{2.53}$$

当该粒子为电子时，(2.53) 式即为描述电子在电磁场运动的 Pauli 方程[2]。也就是说 Dirac 方程描述的是自旋为 1/2 的粒子的运动，φ_1 的两个分量则描述了电子自旋的两个自由度。实际上，通过对 Dirac 方程实施 Foldy-Wouthuysen 变换，我们也能得到 Dirac 方程的非相对论极限即为 Pauli 方程 —— 包含自旋的 Schrödinger 方程。关于 Foldy-Wouthuysen 变换我们将在第 5 章详细讨论。

2.3 Dirac 方程的两个严格解

2.3.1 自由电子

Dirac 方程有两个重要的严格解：一个是自由电子，一个是氢原子。我们首先讨论自由电子。令

$$\varPsi(\boldsymbol{r},t)=\varPsi(\boldsymbol{r})\exp\left(-\frac{\mathrm{i}}{\hbar}Et\right) \tag{2.54}$$

代入自由电子的 Dirac 方程 (2.30) 式，得 $\varPsi(\boldsymbol{r})$ 满足定态 Dirac 方程

$$\left(c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}+\hat{\beta}m_0c^2\right)\varPsi(\boldsymbol{r})=E\varPsi(\boldsymbol{r}) \tag{2.55}$$

为方便，将四分量波函数写成如下二分量形式

$$\varPsi(\boldsymbol{r})=\begin{pmatrix}\varPsi_1(\boldsymbol{r})\\ \varPsi_2(\boldsymbol{r})\\ \varPsi_3(\boldsymbol{r})\\ \varPsi_4(\boldsymbol{r})\end{pmatrix}=\begin{pmatrix}\tilde{\varphi_1}\\ \tilde{\varphi_2}\end{pmatrix} \tag{2.56}$$

其中

$$\tilde{\varphi_1}=\begin{pmatrix}\varPsi_1(\boldsymbol{r})\\ \varPsi_2(\boldsymbol{r})\end{pmatrix},\tilde{\varphi_2}=\begin{pmatrix}\varPsi_3(\boldsymbol{r})\\ \varPsi_4(\boldsymbol{r})\end{pmatrix} \tag{2.57}$$

观察自由电子的 Dirac Hamilton 算符 $\hat{h}_{\mathrm{D}}$(见 (2.31) 式)，仅包含 $\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}$ 和 $\hat{\beta}$ 两种算符，其中 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\beta}$ 为常量算符。不难证明，$\hat{h}_{\mathrm{D}}$ 与动量算符 $\hat{\boldsymbol{p}}$ 及螺旋度算符 $\hat{h}$ 对易，所以这三个对易算符必有共同的本征函数。螺旋度算符 $\hat{h}$ 定义为自旋在动量方向上的投影

$$\hat{h}=\hat{\boldsymbol{S}}\cdot\frac{\hat{\boldsymbol{p}}}{p} \tag{2.58}$$

式中，p 表示动量算符本征值的大小，是一个标量。$\hat{\boldsymbol{S}}$ 为自旋算符[2] 的四维形式

$$\hat{\boldsymbol{S}}=\frac{\hbar}{2}\hat{\boldsymbol{\Sigma}}=\frac{\hbar}{2}\begin{pmatrix}\hat{\boldsymbol{\sigma}} & 0\\ 0 & \hat{\boldsymbol{\sigma}}\end{pmatrix} \tag{2.59}$$

上式中四维 Pauli 算符 $\hat{\boldsymbol{\Sigma}}$ 与自旋算符 $\hat{\boldsymbol{S}}$ 仅相差一个常数 $\hbar/2$，有时也称 Pauli 算符为自旋算符。下面我们就来求 $\hat{h}_{\mathrm{D}}$, $\hat{\boldsymbol{p}}$, $\hat{h}$ 三个对易算符的共同本征函数。

动量算符 $\hat{\boldsymbol{p}}$ 的本征方程为关于变量 x,y,z 的一阶偏微分方程，容易求得其本征函数

$$\Psi_p(\boldsymbol{r})=\begin{pmatrix}\varphi_1^0\\ \varphi_2^0\end{pmatrix}\frac{1}{(2\pi\hbar)^{3/2}}\exp\left(\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}\right) \tag{2.60}$$

式中，φ_1^0 和 φ_2^0 为不包含变量 x,y,z 的二维列矢量。

为求得 φ_1^0 和 φ_2^0，令 (2.60) 式的 $\Psi_p(\boldsymbol{r})$ 也是算符 $\hat{h}$ 的本征函数。本征方程 $\hat{h}\Psi_p(\boldsymbol{r})=C\Psi_p(\boldsymbol{r})$ 可写为

$$\frac{\hbar}{2}\frac{1}{p}\begin{pmatrix}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}} & 0\\ 0 & \hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\end{pmatrix}\begin{pmatrix}\varphi_1^0\\ \varphi_2^0\end{pmatrix}=\pm\frac{\hbar}{2}\begin{pmatrix}\varphi_1^0\\ \varphi_2^0\end{pmatrix} \tag{2.61}$$

上式右边的本征值 $\pm\dfrac{\hbar}{2}$ 是由本征方程的久期行列式定出的。事实上，由自旋的属性我们不难判断自旋在任何方向上的投影均为 $\pm\dfrac{\hbar}{2}$。由 (2.61) 式可知，φ_1^0 和 φ_2^0 都是 $\dfrac{1}{p}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}$ 的本征函数，且本征值相同。因此 φ_1^0 和 φ_2^0 只能相差一个常数 (用 a' 表示)，即 $\varphi_2^0=a'\varphi_1^0$。

现在我们来求 φ_1^0，令

$$\varphi_1^0=\begin{pmatrix}\Psi_1^0\\ \Psi_2^0\end{pmatrix} \tag{2.62}$$

则有

$$\frac{1}{p}\begin{pmatrix}p_z & p_x-\mathrm{i}p_y\\ p_x+\mathrm{i}p_y & -p_z\end{pmatrix}\begin{pmatrix}\Psi_1^0\\ \Psi_2^0\end{pmatrix}=\pm\begin{pmatrix}\Psi_1^0\\ \Psi_2^0\end{pmatrix} \tag{2.63}$$

采用球坐标 (r,ϑ,φ) 来表达动量 $\boldsymbol{p}$，有 $p_x=p\sin\vartheta\cos\varphi$, $p_y=p\sin\vartheta\sin\varphi$, $p_z=p\cos\vartheta$。代入 (2.63) 式，得

$$\begin{pmatrix}\cos\vartheta & \sin\vartheta\mathrm{e}^{-\mathrm{i}\varphi}\\ \sin\vartheta\mathrm{e}^{\mathrm{i}\varphi} & -\cos\vartheta\end{pmatrix}\begin{pmatrix}\Psi_1^0\\ \Psi_2^0\end{pmatrix}=\pm\begin{pmatrix}\Psi_1^0\\ \Psi_2^0\end{pmatrix} \tag{2.64}$$

方程的解为

正本征值　　$\Psi_2^0=\tan\dfrac{\vartheta}{2}\mathrm{e}^{\mathrm{i}\varphi}\Psi_1^0$

负本征值　$\Psi_2^0 = -\cot\dfrac{\vartheta}{2}\mathrm{e}^{\mathrm{i}\varphi}\Psi_1^0$

代回 (2.62) 式，得

正本征值

$$\varphi_1^0 = \varphi_+^0 = \begin{pmatrix} \cos\dfrac{\vartheta}{2}\mathrm{e}^{-\mathrm{i}\frac{\varphi}{2}} \\ \sin\dfrac{\vartheta}{2}\mathrm{e}^{\mathrm{i}\frac{\varphi}{2}} \end{pmatrix} \tag{2.65}$$

负本征值

$$\varphi_1^0 = \varphi_-^0 = \begin{pmatrix} -\sin\dfrac{\vartheta}{2}\mathrm{e}^{-\mathrm{i}\frac{\varphi}{2}} \\ \cos\dfrac{\vartheta}{2}\mathrm{e}^{\mathrm{i}\frac{\varphi}{2}} \end{pmatrix} \tag{2.66}$$

这里把 φ_+^0 和 φ_-^0 写成归一化的对称形式。把此二式代回到 (2.60) 式中，记住 φ_1^0 和 φ_2^0 只差一个常数，得

$$\Psi_{p\pm}(\boldsymbol{r}) = \begin{pmatrix} \varphi_\pm^0 \\ a'_\pm\varphi_\pm^0 \end{pmatrix} \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}\right) \tag{2.67}$$

这是 $\hat{\boldsymbol{p}}$ 和 $\hat{h}$ 的共同本征函数，加号对应的本征值是 $\boldsymbol{p},+\dfrac{\hbar}{2}$，减号对应的本征值是 $\boldsymbol{p},-\dfrac{\hbar}{2}$。上式中的 $a'_\pm$ 是常数。为定此常数，我们把 (2.67) 式代入定态 Dirac 方程 (2.55) 式，得

$$\left(c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}} + \hat{\beta}m_0c^2\right)\begin{pmatrix} \varphi_\pm^0 \\ a'_\pm\varphi_\pm^0 \end{pmatrix} = E'\begin{pmatrix} \varphi_\pm^0 \\ a'_\pm\varphi_\pm^0 \end{pmatrix} \tag{2.68}$$

E' 是 Hamilton 算符的本征值即自由电子的能量。注意到 (2.61) 式，(2.68) 式可改写为

$$\begin{pmatrix} m_0c^2 - E' & \pm cp \\ \pm cp & -m_0c^2 - E' \end{pmatrix}\begin{pmatrix} \varphi_\pm^0 \\ a'_\pm\varphi_\pm^0 \end{pmatrix} = 0 \tag{2.69}$$

方程组中 $\varphi_\pm^0$ 有非零解的条件是系数矩阵的行列式为零，即

$$\begin{vmatrix} m_0c^2 - E' & \pm cp \\ \pm cp & -m_0c^2 - E' \end{vmatrix} = 0 \tag{2.70}$$

解得自由电子的能量　$E' = \pm\sqrt{c^2p^2 + m_0^2c^4} = \pm E$

而 $a'_\pm$ 的解为

$$a'_\pm = \pm a = \pm\frac{E - m_0c^2}{cp} = \pm\frac{cp}{E + m_0c^2} \tag{2.71}$$

将此 $a'_\pm$ 代入 (2.67) 式，并注意到 (2.54) 式，最后得到 $\hat{h}_{\mathrm{D}}, \hat{\boldsymbol{p}}$ 和 $\hat{h}$ 的本征值和共同本征函数，如表 2.1 所列。

表 2.1 自由电子 Dirac 算符 $\hat{h}_D$、动量算符 $\hat{\boldsymbol{p}}$、螺旋度算符 $\hat{h}$ 的本征值和本征函数 $\Psi(\boldsymbol{r},t)$

$\hat{h}_D$	$\hat{\boldsymbol{p}}$	$\hat{h}$	$\Psi(\boldsymbol{r},t)$
E	$\boldsymbol{p}$	$+\dfrac{\hbar}{2}$	$\Phi_{+p+}(\boldsymbol{r},t)=N\begin{pmatrix}\varphi_+^0\\ a\varphi_+^0\end{pmatrix}\exp\left[\dfrac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)\right]$
E	$\boldsymbol{p}$	$-\dfrac{\hbar}{2}$	$\Phi_{+p-}(\boldsymbol{r},t)=N\begin{pmatrix}\varphi_-^0\\ -a\varphi_-^0\end{pmatrix}\exp\left[\dfrac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}-Et)\right]$
$-E$	$\boldsymbol{p}$	$+\dfrac{\hbar}{2}$	$\Phi_{-p+}(\boldsymbol{r},t)=N\begin{pmatrix}a\varphi_+^0\\ -\varphi_+^0\end{pmatrix}\exp\left[\dfrac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}+Et)\right]$
$-E$	$\boldsymbol{p}$	$-\dfrac{\hbar}{2}$	$\Phi_{-p-}(\boldsymbol{r},t)=N\begin{pmatrix}a\varphi_-^0\\ \varphi_-^0\end{pmatrix}\exp\left[\dfrac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}+Et)\right]$

注：φ_+^0 和 φ_-^0 分别见 (2.65) 式和 (2.66) 式；a 见 (2.71) 式；N 见 (2.72) 式。

表 2.1 中的归一化系数 N 为

$$N=\frac{1}{(2\pi\hbar)^{3/2}}\sqrt{\frac{E+m_0c^2}{2E}} \tag{2.72}$$

与 Klein-Gordon 方程一样 (见 2.1 节)，这里我们也得到了负能态解。在非相对论经典力学中一切粒子的总能量都是正的，总能量为负的粒子是不存在的。为解释负能态问题，Dirac 提出了空穴理论，根据该理论，负能态是存在的，但已充满电子。由于 Pauli 不相容原理，正能态的电子不会再落入到负能态去；而负能态的电子海是不能被观察到的。如果在负能态的电子海中出现一个空穴，则能观察到一个正能态的带正电荷的粒子，这就从理论上预言了正电子的存在。果然，1932 年美国物理学家 Anderson 在宇宙射线中首次发现了正电子。

尽管如此，空穴理论只是一个过渡性的理论。从质量的角度，弥漫空间的负电荷却无法观察到是无法接受的。同样在 Dirac 奠基的量子电动力学中，并不需要空穴和电子海的概念，它是关于电子和正电子的圆满的、全面的理论。有兴趣的读者可以进一步阅读 Greiner 和 Reinhardt 编写的 *Quantum Electrodynamics*[3]。

观察表 2.1 中本征函数的两个分量 $(\varphi_+^0,\varphi_-^0)$，得二者模的比值为

$$r=a^2=\frac{c^2p^2}{(E+m_0c^2)^2}\quad (E>0) \tag{2.73}$$

当电子的速度 v 远小于光速 c 时，由 (1.14) 式、(1.33) 式和 (1.42) 式，得

$$p=\frac{m_0}{\sqrt{1-\dfrac{v^2}{c^2}}}v\approx m_0v$$

$$E=\sqrt{c^2p^2+m_0^2c^4}\approx m_0c^2 \tag{2.74}$$

代入 (2.73) 式，得

$$r \approx \frac{v^2}{4c^2} \ll 1 \tag{2.75}$$

即当电子的速度 v 远小于光速 c 时，上分量的模远大于下分量的模，所以波函数的上下两个分量又称为大小分量波函数，分别用 Ψ_{L}, Ψ_{S} 表示。对正电子而言 $(E<0)$，结论正好和电子相反，正电子波函数上分量的模远小于下分量的模。由于 Dirac 方程解的这个特性，催生了二分量求解方法，即忽略小分量波函数的贡献，将四分量的 Dirac 方程简化为只有两个分量的方程再求解。我们将在第 5 章具体介绍二分量方法。

综上所述，在相对论理论中，自由电子的态函数与非相对论理论相同，仍是平面波，在新的 4 维空间中表示为一列矩阵。当动量的本征值取 z 轴方向，这个一列矩阵成为

$$\begin{pmatrix}1\\0\\a\\0\end{pmatrix},\quad \begin{pmatrix}0\\1\\0\\-a\end{pmatrix},\quad \begin{pmatrix}-a\\0\\1\\0\end{pmatrix},\quad \begin{pmatrix}0\\a\\0\\1\end{pmatrix} \tag{2.76}$$

它们表示能量 E 的符号和自旋 $\hat{S}_z$ 的方向。由 (2.73) 式、(2.75) 式可知，在低能极限时 a 与 1 相比是一个小量。由此可以看出自旋空间的四个基矢为

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\quad \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \tag{2.77}$$

前两个描述正能态，后两个描述负能态。对于自由电子的态函数 (2.76) 式，$E'=E$ 的状态中有少量负能态叠加在其上，同样 $E'=-E$ 的状态中也有少量正能态的成分。

2.3.2 氢原子

由 (2.43) 式，可得氢原子的定态 Dirac 方程是

$$\left(c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}+\hat{\beta}m_0c^2-\frac{e_1^2}{r}\right)\Psi(\boldsymbol{r})=E\Psi(\boldsymbol{r}) \tag{2.78}$$

上式左边括号内的算符之和即为氢原子 Dirac Hamilton 算符 $\hat{H}$，式中 $e_1=e/\sqrt{4\pi\epsilon_0}$，$\epsilon_0$ 为真空介电常量。

与求解自由电子 Dirac 方程的步骤类似，我们需要找到一组与氢原子 Dirac Hamilton 算符对易的算符，再求这些算符共同的本征函数。容易证明，$\hat{\boldsymbol{J}}^{\boldsymbol{2}}$, $\hat{J}_z$, $\hbar\hat{K}$

三个算符均满足与氢原子 Dirac Hamilton 算符对易的要求。其中 $\hat{\boldsymbol{J}}$, $\hat{J}_z$ 为大家熟悉的总角动量算符及总角动量算符在 z 轴的分量。

$$\hat{\boldsymbol{J}}^2 = (\hat{\boldsymbol{L}} + \hat{\boldsymbol{S}})^2 \tag{2.79}$$

$$\hat{J}_z = \hat{L}_z + \hat{S}_z \tag{2.80}$$

式中，$\hat{\boldsymbol{L}}$, $\hat{\boldsymbol{S}}$ 分别为轨道角动量和自旋角动量算符。算符 $\hbar\hat{K}$ 定义为

$$\hbar\hat{K} = \beta(\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\Sigma}} + \hbar) \tag{2.81}$$

式中，$\hat{\boldsymbol{\Sigma}}$ 为四维 Pauli 算符 (见 (2.59) 式)。显然，算符 $\hat{\boldsymbol{J}}$, $\hat{J}_z$ 和 $\hbar\hat{K}$ 都是块对角形式

$$\hat{\boldsymbol{J}}^2 = \begin{pmatrix} \left(\hat{\boldsymbol{L}} + \dfrac{\hbar}{2}\hat{\boldsymbol{\sigma}}\right)^2 & 0 \\ 0 & \left(\hat{\boldsymbol{L}} + \dfrac{\hbar}{2}\hat{\boldsymbol{\sigma}}\right)^2 \end{pmatrix}, \quad \hat{J}_z = \begin{pmatrix} \hat{L}_z + \dfrac{\hbar}{2}\hat{\sigma}_z & 0 \\ 0 & \hat{L}_z + \dfrac{\hbar}{2}\hat{\sigma}_z \end{pmatrix} \tag{2.82}$$

$$\hbar\hat{K} = \begin{pmatrix} \hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\sigma}} + \hbar & 0 \\ 0 & -\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{\sigma}} - \hbar \end{pmatrix} \tag{2.83}$$

以上矩阵的矩阵元都是 2×2 矩阵，所以以上三个算符均为 4×4 矩阵。

由于 $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$, $\hbar\hat{K}$ 均和氢原子 Dirac Hamilton 算符 $\hat{H}$ 对易，则这四个算符必有共同的本征函数，用 $\Psi_{jmk}(\boldsymbol{r})$ 表示。与求解自由电子 Dirac 方程时的方法类似，我们将四维本征函数 $\Psi_{jmk}(\boldsymbol{r})$ 写成两个二维矩阵的一列形式

$$\Psi_{jmk}(\boldsymbol{r}) = \begin{pmatrix} \Psi_a(\boldsymbol{r}) \\ \Psi_b(\boldsymbol{r}) \\ \Psi_c(\boldsymbol{r}) \\ \Psi_d(\boldsymbol{r}) \end{pmatrix} = \begin{pmatrix} \tilde{\varphi}^1_{jmk} \\ \tilde{\varphi}^2_{jmk} \end{pmatrix} \tag{2.84}$$

$\Psi_{jmk}(\boldsymbol{r})$ 满足

$$\begin{aligned} \hat{\boldsymbol{J}}^2 \Psi_{jmk}(\boldsymbol{r}) &= j(j+1)\hbar^2 \Psi_{jmk}(\boldsymbol{r}) \\ \hat{J}_z \Psi_{jmk}(\boldsymbol{r}) &= m_j\hbar \Psi_{jmk}(\boldsymbol{r}) \\ \hbar\hat{K} \Psi_{jmk}(\boldsymbol{r}) &= \hbar k \Psi_{jmk}(\boldsymbol{r}) \end{aligned} \tag{2.85}$$

式中，$j = 1/2,\ 3/2,\ 5/2,\ \cdots; m_j = -j, -j+1, \cdots, j$; k 值待定。将 (2.82) 式、(2.83) 式和 (2.84) 式代入上式，得

$$\left(\hat{\boldsymbol{L}}+\frac{\hbar}{2}\hat{\boldsymbol{\sigma}}\right)^2\tilde{\varphi}^{1,2}_{jmk}=j(j+1)\hbar^2\tilde{\varphi}^{1,2}_{jmk}$$
$$\left(\hat{L}_z+\frac{\hbar}{2}\hat{\sigma}_z\right)\tilde{\varphi}^{1,2}_{jmk}=m_j\hbar\tilde{\varphi}^{1,2}_{jmk}$$
$$\left(\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{\sigma}}+\hbar\right)\tilde{\varphi}^{1,2}_{jmk}=\pm k\hbar\tilde{\varphi}^{1,2}_{jmk} \tag{2.86}$$

第三式中 $\tilde{\varphi}^1_{jmk}$ 与本征值 $k\hbar$ 对应，$\tilde{\varphi}^2_{jmk}$ 与本征值 $-k\hbar$ 对应。将 (2.86) 式中第一式左边的平方算符展开，利用第三式的结果及

$$\hat{\boldsymbol{\sigma}}^2=3\boldsymbol{I_2}=\begin{pmatrix}3&0\\0&3\end{pmatrix} \tag{2.87}$$

有

$$\begin{aligned}\hat{\boldsymbol{L}}^2\tilde{\varphi}^{1,2}_{jmk}&=\left[j\left(j+1\right)\hbar^2-\frac{3\hbar^2}{4}-\hbar\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{\sigma}}\right]\tilde{\varphi}^{1,2}_{jmk}\\&=\left[j\left(j+1\right)\hbar^2-\frac{3\hbar^2}{4}-\left(\pm k-1\right)\hbar^2\right]\tilde{\varphi}^{1,2}_{jmk}\end{aligned} \tag{2.88}$$

由算符 $\hbar\hat{K}$ 的定义 (2.81) 式，并利用 (2.51) 式和 (2.79) 式，得

$$\begin{aligned}\left(\hbar\hat{K}\right)^2&=\left(\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{\Sigma}}+\hbar\boldsymbol{I}_4\right)^2\\&=\hat{\boldsymbol{J}}^{\boldsymbol{2}}+\frac{1}{4}\hbar^2\boldsymbol{I}_4\end{aligned} \tag{2.89}$$

$\boldsymbol{I}_4$ 是 4×4 单位矩阵。所以

$$k^2\hbar^2=j(j+1)\hbar^2+\frac{1}{4}\hbar^2 \tag{2.90}$$

代入 (2.88) 式，得

$$\hat{\boldsymbol{L}}^2\tilde{\varphi}^{1,2}_{jmk}=k\left(k\mp1\right)\hbar^2\tilde{\varphi}^{1,2}_{jmk} \tag{2.91}$$

由上式可知 $\tilde{\varphi}^{1,2}_{jmk}$ 是 $\hat{\boldsymbol{L^2}}$ 的本征函数，本征值为 $k(k\mp1)\hbar^2$，其中 $\tilde{\varphi}^1_{jmk}$ 与本征值 $k(k-1)$ 对应，$\tilde{\varphi}^2_{jmk}$ 与本征值 $k(k+1)$ 对应。

已知轨道角动量算符 $\hat{\boldsymbol{L}}^2$ 的本征值为 $l(l+1)\hbar^2$，l 表示轨道角量子数，与 (2.91) 式比较得 l 和 k 的换算关系

$$l^{1,2}(l^{1,2}+1)=k(k\mp1) \tag{2.92}$$

接下来我们将 $\hat{J}_z=\hat{L}_z+\dfrac{\hbar}{2}\hat{\sigma}_z$ 和 (2.84) 式代入 (2.86) 式中的第二式，得

$$\hat{L}_z\Psi_{a,c}=\left(m_j-\frac{1}{2}\right)\hbar\Psi_{a,c}$$
$$\hat{L}_z\Psi_{b,d}=\left(m_j+\frac{1}{2}\right)\hbar\Psi_{b,d} \tag{2.93}$$

联合 (2.91) 式，可知 $\Psi_{a,b,c,d}$ 是 $\hat{\boldsymbol{L}}^2$ 和 $\hat{L}_z$ 的共同本征函数，可用球谐函数 $Y_{l,m}$ 来表示

$$\begin{aligned}\Psi_{a,c} &= f'(r)Y_{l^{1,2},m_j-\frac{1}{2}}\\ \Psi_{b,d} &= g'(r)Y_{l^{1,2},m_j+\frac{1}{2}}\end{aligned} \tag{2.94}$$

将上式代入 $\tilde{\varphi}^{1,2}_{jmk}$，再代入 (2.86) 式中的第三式，我们发现 $f'(r)$ 和 $g'(r)$ 只相差一个常数因子

$$\frac{f'(r)}{g'(r)} = \frac{-\sqrt{\left(l^{1,2}+m_j+\frac{1}{2}\right)\left(l^{1,2}-m_j+\frac{1}{2}\right)}}{m_j+\frac{1}{2}\mp k} \tag{2.95}$$

由 (2.90) 式，有

$$k = \pm\left(j+\frac{1}{2}\right) \tag{2.96}$$

联合 (2.92) 式，得

当 $k = +\left(j+\frac{1}{2}\right) > 0$ 时

$$\begin{aligned}l^1 &= k-1 = j-\frac{1}{2}\\ l^2 &= k = j+\frac{1}{2}\end{aligned} \tag{2.97}$$

当 $k = -\left(j+\frac{1}{2}\right) < 0$ 时

$$\begin{aligned}l^1 &= -k = j+\frac{1}{2}\\ l^2 &= -(k+1) = j-\frac{1}{2}\end{aligned} \tag{2.98}$$

将上面的 $l^{1,2}$ 代入 (2.95) 式，再代回到 (2.84) 式，最后得 $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$, $\hbar\hat{K}$ 的共同本征函数 Ψ_{jmk}。

$$\Psi_{jmk} = \begin{pmatrix}\tilde{\varphi}^1_{jmk}\\ \tilde{\varphi}^2_{jmk}\end{pmatrix} \tag{2.99}$$

当 $k>0$ 时

$$\begin{pmatrix}\tilde{\varphi}^1_{jmk}\\ \tilde{\varphi}^2_{jmk}\end{pmatrix} = \begin{pmatrix}f(r)\chi^1_{jm}\\ g(r)\chi^2_{jm}\end{pmatrix} \tag{2.100}$$

式中，m 是 m_j 的简写；$f(r)$ 和 $g(r)$ 是待定函数；χ^1_{jm}，χ^2_{jm} 称为角动量二维旋量

$$\chi_{jm}^1=\begin{pmatrix}\sqrt{\dfrac{j+m}{2j}}Y_{j-\frac{1}{2},m-\frac{1}{2}}\\ \sqrt{\dfrac{j-m}{2j}}Y_{j-\frac{1}{2},m+\frac{1}{2}}\end{pmatrix},\quad \chi_{jm}^2=\begin{pmatrix}-\sqrt{\dfrac{j-m+1}{2(j+1)}}Y_{j+\frac{1}{2},m-\frac{1}{2}}\\ \sqrt{\dfrac{j+m+1}{2(j+1)}}Y_{j+\frac{1}{2},m+\frac{1}{2}}\end{pmatrix} \tag{2.101}$$

当 $k<0$ 时

$$\begin{pmatrix}\tilde{\varphi}_{jmk}^1\\ \tilde{\varphi}_{jmk}^2\end{pmatrix}=\begin{pmatrix}f(r)\chi_{jm}^2\\ g(r)\chi_{jm}^1\end{pmatrix} \tag{2.102}$$

式中，$j=1/2,\ 3/2,\ 5/2,\cdots;m_j=-j,-j+1,\cdots,j$。

我们习惯上用光谱符号 $\mathrm{s,p,d},\cdots$ 来表示角动量二维旋量 $\chi_{jm}^{1,2}$，如表 2.2 所示。表 2.3 则列出了部分原子标记所对应的角动量二维旋量。

表 2.2　量子数和原子的光谱符号

l	0	1	1	2	2	3	3
j	1/2	1/2	3/2	3/2	5/2	5/2	7/2
k	1	−1	2	−2	3	−3	4
光谱符号	$\mathrm{s}_{1/2}$	$\mathrm{p}_{1/2}$	$\mathrm{p}_{3/2}$	$\mathrm{d}_{3/2}$	$\mathrm{d}_{5/2}$	$\mathrm{f}_{5/2}$	$\mathrm{f}_{7/2}$
	s	p_-	p_+	d_-	d_+	f_-	f_+

表 2.3　原子 s, p 轨道的角动量二维旋量

$\mathrm{s}_{1/2}$	$\chi^1_{\frac{1}{2},\frac{1}{2}}=\begin{pmatrix}Y_{0,0}\\0\end{pmatrix}=\begin{pmatrix}\mathrm{s}\\0\end{pmatrix}$
	$\chi^1_{\frac{1}{2},-\frac{1}{2}}=\begin{pmatrix}0\\Y_{0,0}\end{pmatrix}=\begin{pmatrix}0\\\mathrm{s}\end{pmatrix}$
$\sqrt{3}\mathrm{p}_{1/2}$	$\chi^2_{\frac{1}{2},\frac{1}{2}}=\begin{pmatrix}-Y_{1,0}\\\sqrt{2}Y_{1,1}\end{pmatrix}=\begin{pmatrix}-\mathrm{p}_0\\\sqrt{2}\mathrm{p}_1\end{pmatrix}$
	$\chi^2_{\frac{1}{2},-\frac{1}{2}}=\begin{pmatrix}-\sqrt{2}Y_{1,1}\\Y_{1,0}\end{pmatrix}=\begin{pmatrix}-\sqrt{2}\mathrm{p}_1\\\mathrm{p}_0\end{pmatrix}$
$\sqrt{3}\mathrm{p}_{3/2}$	$\chi^1_{\frac{3}{2},\frac{3}{2}}=\begin{pmatrix}\frac{1}{\sqrt{3}}Y_{1,1}\\0\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{3}}\mathrm{p}_1\\0\end{pmatrix}$
	$\chi^1_{\frac{3}{2},\frac{1}{2}}=\begin{pmatrix}\sqrt{2}Y_{1,0}\\Y_{1,1}\end{pmatrix}=\begin{pmatrix}\sqrt{2}\mathrm{p}_0\\\mathrm{p}_1\end{pmatrix}$
	$\chi^1_{\frac{3}{2},-\frac{1}{2}}=\begin{pmatrix}Y_{1,-1}\\\sqrt{2}Y_{1,0}\end{pmatrix}=\begin{pmatrix}\mathrm{p}_{-1}\\\sqrt{2}\mathrm{p}_0\end{pmatrix}$
	$\chi^1_{\frac{3}{2},-\frac{3}{2}}=\begin{pmatrix}0\\\frac{1}{\sqrt{3}}Y_{1,-1}\end{pmatrix}=\begin{pmatrix}0\\\frac{1}{\sqrt{3}}\mathrm{p}_{-1}\end{pmatrix}$

2.3.3　径向方程的解

已经求得的 Ψ_{jmk}，即 (2.99) 式、(2.100) 式和 (2.102) 式是 $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$, $\hbar\hat{K}$ 的本征函数，其中只有 $f(r)$ 和 $g(r)$ 待定。现在令 Ψ_{jmk} 是 $\hat{H}$ 的本征函数，由此定出 $f(r)$ 和 $g(r)$ 两个函数，以及 $\hat{H}$ 的本征值 E。

由附录 B 中的 (B.13) 式，有

$$\boldsymbol{r}\times(\boldsymbol{r}\times\nabla)=\boldsymbol{r}(\boldsymbol{r}\cdot\nabla)-r^2\nabla \tag{2.103}$$

又 $\boldsymbol{r}\cdot\nabla=r\dfrac{\partial}{\partial r}$，$\hat{\boldsymbol{L}}=-\mathrm{i}\hbar(\hat{\boldsymbol{r}}\times\nabla)$，代入 (2.103) 式，整理得

$$\nabla=-\frac{\mathrm{i}\hat{\boldsymbol{r}}\times\hat{\boldsymbol{L}}}{\hbar r^2}+\hat{\boldsymbol{e}}_r\frac{\partial}{\partial r} \tag{2.104}$$

式中，$\hat{\boldsymbol{e}}_r$ 代表径向变量 $\boldsymbol{r}$ 的单位矢量算符。利用上式，$\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}$ 可改写为

$$\begin{aligned}\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}&=-\mathrm{i}\hbar\hat{\boldsymbol{\alpha}}\cdot\nabla\\&=\frac{-\hat{\boldsymbol{\alpha}}\cdot\left(\hat{\boldsymbol{e}}_r\times\hat{\boldsymbol{L}}\right)}{r}-\mathrm{i}\hbar\left(\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{e}}_r\right)\frac{\partial}{\partial r}\end{aligned} \tag{2.105}$$

另外，由 Pauli 矢量规则 (见附录 B 中 (B.3) 式)，任意两个矢量 $\boldsymbol{A}$, $\boldsymbol{B}$ 与 $\boldsymbol{\alpha}$, $\boldsymbol{\Sigma}$ 满足

$$(\boldsymbol{\alpha}\cdot\boldsymbol{A})(\boldsymbol{\Sigma}\cdot\boldsymbol{B})=-\mathrm{i}\alpha_x\alpha_y\alpha_z\boldsymbol{A}\cdot\boldsymbol{B}+\mathrm{i}\boldsymbol{\alpha}\cdot\boldsymbol{A}\times\boldsymbol{B} \tag{2.106}$$

式中，$\alpha_i(i=x,y,z)$ 为 $\boldsymbol{\alpha}$ 在 x,y,z 轴的分量。利用 (2.106) 式可得

$$\begin{aligned}\mathrm{i}\hat{\boldsymbol{\alpha}}\cdot\left(\hat{\boldsymbol{e}}_r\times\hat{\boldsymbol{L}}\right)&=\left(\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{e}}_r\right)\left(\hat{\boldsymbol{\Sigma}}\cdot\hat{\boldsymbol{L}}\right)\\&=\left(\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{e}}_r\right)\left(\beta\hbar\hat{K}-\hbar\right)\end{aligned} \tag{2.107}$$

将上式代入 (2.105) 式，再代入 (2.78) 式，得氢原子 Dirac Hamilton 算符

$$\hat{H}=-\mathrm{i}\hbar c\hat{\alpha}_r\left[\frac{\partial}{\partial r}-\frac{1}{r}\left(\beta\hat{K}-1\right)\right]+m_0c^2\beta-\frac{e_1^2}{r} \tag{2.108}$$

式中，$\hat{\alpha}_r=\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{e}}_r$。将 (2.99) 式连同 (2.101) 式代入 $\hat{H}\Psi=E\Psi$ 中，并注意到 (2.99) 式是 $\hbar\hat{K}$ 的本征函数，得两个 2×2 方程

$$\begin{aligned}\left(m_0c^2-\frac{e_1^2}{r}-E\right)f(r)\chi^1&=-\left[-\mathrm{i}\hbar c\frac{\partial}{\partial r}-\frac{\mathrm{i}c\hbar}{r}(k+1)\right]g(r)\hat{\sigma}_r\chi^2\\\left(m_0c^2+\frac{e_1^2}{r}+E\right)g(r)\chi^2&=\left[-\mathrm{i}\hbar c\frac{\partial}{\partial r}+\frac{\mathrm{i}c\hbar}{r}(k-1)\right]f(r)\hat{\sigma}_r\chi^1\end{aligned} \tag{2.109}$$

比较等式左右两边，得角动量二维旋量 χ^1 和 χ^2 满足

$$\begin{aligned}\chi^1&=\hat{\sigma}_r\chi^2\\\chi^2&=\hat{\sigma}_r\chi^1\end{aligned} \tag{2.110}$$

径向函数 $f(r)$ 和 $g(r)$ 则满足以下方程

$$\begin{aligned}&\left(m_0c^2-\frac{e_1^2}{r}-E\right)f(r)-\mathrm{i}\hbar c\frac{\mathrm{d}g(r)}{\mathrm{d}r}-\frac{\mathrm{i}c\hbar}{r}(k+1)\,g(r)=0\\&\left(-m_0c^2-\frac{e_1^2}{r}-E\right)g(r)-\mathrm{i}\hbar c\frac{\mathrm{d}f(r)}{\mathrm{d}r}+\frac{\mathrm{i}c\hbar}{r}(k-1)\,f(r)=0\end{aligned}\tag{2.111}$$

这是两个联立的微分方程，取

$$F(r)=rf(r),\quad \mathrm{i}G(r)=rg(r)\tag{2.112}$$

则 $F(r)$ 和 $G(r)$ 的方程成为 (令 $\epsilon=E/m_0c^2$)

$$\begin{aligned}&\frac{\mathrm{d}G}{\mathrm{d}r}+\frac{k}{r}G+\left[\frac{m_0c}{\hbar}(1-\epsilon)-\frac{\alpha}{r}\right]F=0\\&\frac{\mathrm{d}F}{\mathrm{d}r}-\frac{k}{r}F+\left[\frac{m_0c}{\hbar}(1+\epsilon)+\frac{\alpha}{r}\right]G=0\end{aligned}\tag{2.113}$$

式中，α 是精细结构常数。

$$\alpha=\frac{e_1^2}{\hbar c}=\frac{q^2}{4\pi\epsilon_0\hbar c}\tag{2.114}$$

如果采用原子单位 (atomic unit，a.u.，详见附录 C)，即分别取电子的静止质量 (m_0)、基本电荷 (e)、约化 Planck 常量 ($\hbar$)、Coulomb 常数 ($1/(4\pi\epsilon_0)$) 作为质量、电荷、角动量和 Coulomb 力的单位，有

$$\begin{aligned}m_0=e=\hbar=\frac{1}{4\pi\epsilon_0}=1\\\alpha=\frac{1}{c}\approx\frac{1}{137}\end{aligned}\tag{2.115}$$

下面求 $F(r)$ 和 $G(r)$ 满足平方可积条件下的解。为此，先求方程 (2.113) 式，当 $r\to\infty$ 时的渐近解 $F'(r)$, $G'(r)$，然后在 $F'(r)$ 和 $G'(r)$ 上各乘以一个待求函数为 (2.113) 式的正式解，最后用级数法去求那两个待求函数。

当 $r\to\infty$ 时，方程 (2.113) 式成为

$$\begin{aligned}&\frac{\mathrm{d}G'}{\mathrm{d}r}=-F'\left[\frac{m_0c}{\hbar}(1-\epsilon)\right]\\&\frac{\mathrm{d}F'}{\mathrm{d}r}=G'\left[\frac{-m_0c}{\hbar}(1+\epsilon)\right]\end{aligned}\tag{2.116}$$

它们的平方可积解可写为

$$F'=C_1\mathrm{e}^{-\frac{\rho}{2}},\qquad G'=C_2\mathrm{e}^{-\frac{\rho}{2}}\tag{2.117}$$

式中

$$\rho=\frac{2m_0c}{\hbar}\sqrt{1-\epsilon^2}\,r \tag{2.118}$$

有了渐近解后，设 (2.113) 式的正式解 $F(r)$ 和 $G(r)$ 为

$$F=F'F'',\quad G=G'G'' \tag{2.119}$$

为方便计算，设 $u=F''+G'',\nu=F''-G''$，而将 (2.119) 式写成

$$F=\sqrt{1+\epsilon}\mathrm{e}^{-\frac{\rho}{2}}(u+\nu),\quad G=\sqrt{1-\epsilon}\mathrm{e}^{-\frac{\rho}{2}}(u-\nu) \tag{2.120}$$

将 F 和 G 代回 (2.113) 式，并将自变量由 r 改为无量纲的 ρ(见 (2.118) 式)，得 u 和 ν 的方程为

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}\rho}&=\left(\rho-\frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}}\right)\frac{u}{\rho}+\left(k-\frac{\alpha}{\sqrt{1-\epsilon^2}}\right)\frac{\nu}{\rho}\\ \frac{\mathrm{d}\nu}{\mathrm{d}\rho}&=\frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}}\frac{\nu}{\rho}+\left(k+\frac{\alpha}{\sqrt{1-\epsilon^2}}\right)\frac{u}{\rho}\end{aligned} \tag{2.121}$$

用级数解法，令

$$u=\rho^{\lambda}\sum_{\mu=0}^{\infty}a_{\mu}\rho^{\mu},\quad \nu=\rho^{\lambda}\sum_{\mu=0}^{\infty}b_{\mu}\rho^{\mu} \tag{2.122}$$

式中，ρ^{λ} 是两个级数的最低幂，即 a_0 和 b_0 至少有一个不为零。将上式代入 (2.121) 式，得 a_{μ} 与 b_{μ} 的方程

$$\begin{aligned}\left(\lambda+\mu+\frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}}\right)a_{\mu}-a_{\mu-1}&=\left(k-\frac{\alpha}{\sqrt{1-\epsilon^2}}\right)b_{\mu}\\ \left(\lambda+\mu-\frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}}\right)b_{\mu}&=\left(k+\frac{\alpha}{\sqrt{1-\epsilon^2}}\right)a_{\mu}\end{aligned} \tag{2.123}$$

现在由此式求 λ 及 a_{μ}, b_{μ}。首先取 $\mu=0$，则根据规定 $a_{\mu-1}=0$，由上式得 a_0 与 b_0 的方程

$$\begin{aligned}\left(\lambda+\frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}}\right)a_0&=\left(k-\frac{\alpha}{\sqrt{1-\epsilon^2}}\right)b_0\\ \left(\lambda-\frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}}\right)b_0&=\left(k+\frac{\alpha}{\sqrt{1-\epsilon^2}}\right)a_0\end{aligned} \tag{2.124}$$

上式有非零解的条件是其系数行列式为零，由此得 λ 为

$$\lambda=\pm\sqrt{k^2-\alpha^2} \tag{2.125}$$

但是，在径向函数的归一化条件 (注意到 (2.112) 式)

$$\int\left(|F|^2+|G|^2\right)\mathrm{d}r=1 \tag{2.126}$$

中，当 $r \to 0$ 即 $\rho \to 0$ 时，被积函数趋于 $\rho^{2\lambda}$，为使被积函数平方可积，必须取 $\lambda > -1/2$，而 $|k|$ 的最小值为 1，所以 λ 在 (2.125) 式中只能取正值。关于 λ 只能取正值的详细推导读者可参见 Greiner 编写的 *Relativistic Quantum Mechanics: Wave Equations* 一书[1]。

由 (2.124) 式和 (2.123) 式的第二式分别求得

$$\frac{a_0}{b_0} = \frac{-n'}{k + \dfrac{\alpha}{\sqrt{1-\epsilon^2}}}, \quad \frac{a_\mu}{b_\mu} = \frac{n'-\mu}{k + \dfrac{\alpha}{\sqrt{1-\epsilon^2}}} \tag{2.127}$$

式中，n' 为

$$n' = -\lambda + \frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}} \tag{2.128}$$

将 a_μ/b_μ 代入 (2.123) 式的第一式，得系数 a_μ 和 b_μ 的递推公式

$$\begin{aligned} a_\mu &= -\frac{n'-\mu}{\mu(2\lambda+\mu)} a_{\mu-1} \\ b_\mu &= -\frac{n'-\mu+1}{\mu(2\lambda+\mu)} b_{\mu-1} \end{aligned} \tag{2.129}$$

或

$$\begin{aligned} a_{\mu+1} &= -\frac{n'-\mu-1}{(\mu+1)(2\lambda+\mu+1)} a_\mu \\ b_{\mu+1} &= -\frac{n'-\mu}{(\mu+1)(2\lambda+\mu+1)} b_\mu \end{aligned} \tag{2.130}$$

观察上式我们发现，当 $r \to \infty$ ，即 $\rho \to \infty$ 时，$\mu \gg 1$(见 (2.122) 式)

$$\frac{a_\mu}{a_{\mu-1}} \to \frac{\mu}{\mu^2} = \frac{1}{\mu} \tag{2.131}$$

所以此时 u 和 ν 的级数展开趋于 e^ρ，这使得由 (2.120) 式得到的 F 和 G 不能平方可积。为使 F 和 G 满足平方可积的条件，必须使 u 和 ν 成为多项式而不是无穷级数，即展开系数 a_μ 和 b_μ 在若干个展开项后为零。

由 (2.129) 式，可知递推公式 (2.129) 式中的 n' 必须是整数

$$n' = -\lambda + \frac{\alpha\epsilon}{\sqrt{1-\epsilon^2}} = 0, 1, 2, 3, \cdots \tag{2.132}$$

$n' = 0$ 的情况需另外考虑。当 $n' = 0$ 时，由 (2.130) 式第二式得展开系数 $b_\mu = 0$ $(\mu > 0)$，则由 (2.122) 式知 $\nu = \rho^\lambda b_0$。而 u 则必须当 a_0 为零时才不致成为无穷级数，即有 $a_0 = 0$。这时 (2.127) 式 $a_0/b_0 = 0$，而其右方分子 $n' = 0$，所以分母必须不为零，即

$$k + \frac{\alpha}{\sqrt{1-\epsilon^2}} \neq 0 \tag{2.133}$$

然而，由 (2.125) 式及 (2.128) 式得出，当 $n'=0$ 时恰好有

$$k=\pm\frac{\alpha}{\sqrt{1-\epsilon^2}} \tag{2.134}$$

可见，为满足 (2.133) 式，k 不能取负值。我们于是得出结论，当 $n'=0$ 时，k 只能取正值。

引入新量子数 n

$$n=n'+|k| \quad (n=1,2,3,\cdots) \tag{2.135}$$

这样，描述氢原子量子态的量子数有 n,k,j,m_j4 个；其中 n 的取值范围为 (2.135) 式；$k=\pm(j+1/2)=\pm1,\pm2,\cdots,+n$ (没有 $-n$)；$j=1/2,\ 3/2,\ 5/2,\cdots,n-1/2$；$m_j$ 取 $+j$ 到 $-j$ 之间各值。至于相应的本征函数，为 (2.84) 式，其中所有部分均已算出。氢原子 Dirac Hamilton 算符 (见 (2.78) 式) 的本征值 E，可由 (2.132) 式得出

$$E=m_0c^2\epsilon=\frac{m_0c^2}{\sqrt{1+\dfrac{\alpha^2}{\left(n-|k|+\sqrt{k^2-\alpha^2}\right)^2}}} \tag{2.136}$$

这是氢原子能量的准确值。此式与实验符合得很好。但是，根据此式氢原子的 $2s_{1/2}$ 和 $2p_{1/2}$，以及 $3p_{3/2}$ 和 $3d_{3/2}$ 等具有相同 n 和 $|k|$ 值的态的能量是相同的，而实验上测得它们的能级有微小的差异，如 1947 年 Lamb 和 Retherford 测得 $2s_{1/2}$ 比 $2p_{1/2}$ 高出约 0.004 meV，现在人们称 $2s_{1/2}$ 能级的这一移位为“Lamb 移位” (Lamb-shift) 。根据量子电动力学的解释，原子能级的 Lamb 移位来自电磁场的真空极化和电子的自能过程①。由于这两项对 Hamilton 量的贡献非常小 (约在 α^3 数量级，α 为精细结构常数)，所以在一般化学计算中我们通常不予考虑。

将 (2.136) 式作级数展开，并忽略小项 ($<\alpha^6$)，得近似式为

$$E\approx m_0c^2\left[1-\frac{\alpha^2}{2n^2}-\frac{\alpha^4}{2n^3}\left(\frac{1}{|k|}-\frac{3}{4n}\right)\right] \tag{2.137}$$

上式右边第二项即是非相对论氢原子的能量。至此，以自由电子、氢原子为例我们已经完成了 Dirac 方程的求解。由此可以得到 Dirac 方程解的一些定性结论：

(1) 主量子数 n 和轨道角量子数 l 相同但总角动量 j 不同的能级是自旋–轨道分裂的。

(2) 本征函数是归一化的。

(3) 由于径向函数 f 和 g 在不同的 r 值存在节点 (函数值为零的点)，波函数径向密度 f^2+g^2 不存在节点，如图 2.1 所示。图 2.1 显示的是氢原子 $2s_{1/2}$ 轨道

① 关于真空极化和电子自能的详细解释见附录D

径向函数 $F(r), G(r)$ 及 F^2+G^2 随 r 的变化。为便于显示，图中 $G(r)$ 值被放大了 100 倍。$F(r)$ 和 $G(r)$ 与真实径向函数 $f(r)$ 和 $g(r)$ 的关系见 (2.112) 式。

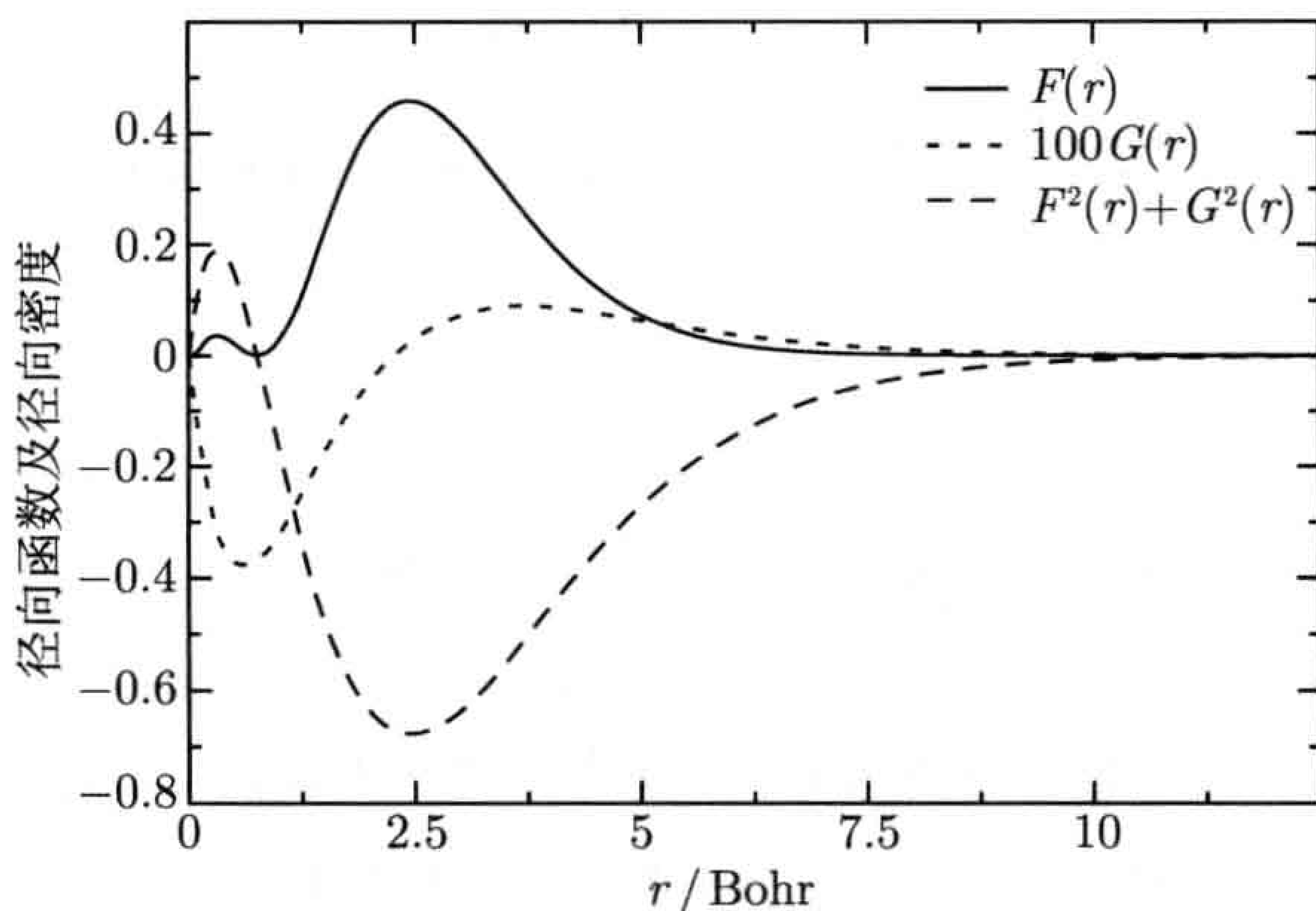


图 2.1　氢原子 $2s_{1/2}$ 轨道径向函数 $F(r), G(r)$ 及径向密度 $F^2(r)+G^2(r)$

(4) 与非相对论 Schrödinger 方程得到的径向电子密度相比，Dirac 方程得到的径向电子密度更紧密，称为相对论收缩，径向电子密度的相对论收缩和自旋–轨道耦合称为直接相对论效应(详见第 7 章)。值得注意的是，在 Dirac 和 Schrödinger 两种图像中径向变量 r 的含义略有不同。在 2.2.4 节我们已经提到对 Dirac 方程进行 Foldy-Wouthuysen 变换 (U_{FW}) 即能得到 Schrödinger 方程，Dirac 方程和 Schrödinger 方程的径向变量 r 的变换关系为[4]

$$r_{\mathrm{Dirac}} = \hat{U}_{\mathrm{FW}} r_{\mathrm{Schrödinger}} \hat{U}_{\mathrm{FW}}^{\dagger} \tag{2.138}$$

对重原子的内层轨道，二者的差异可以非常大，它们对轨道收缩的贡献可以达到 50%[4]。

(5) Dirac 方程的四维波函数解可以写成两个二维旋量的形式，分别称为上分量波函数和下分量波函数。对正能解即 $E>0$，上分量波函数的模远大于下分量波函数的模。从图 2.1 我们也可以看到，下分量波函数 $G(r)$ 的值在放大 100 倍后才能在图中清楚显示。对负能解，正好相反 ——上分量波函数的模远小于下分量波函数的模。

(6) 当 $|k|=1$ 时，波函数在原点存在奇点。由系数递推公式 (2.129) 式，(2.112) 式, (2.120) 式及 (2.122) 式，可得 $|k|=1$ 时的径向函数 $f(r)$ 为

$$f(r) = \frac{1}{r^{|\lambda-1|}}\sqrt{1+\epsilon}\mathrm{e}^{-\frac{1}{2}\rho}\left(2m_0c\sqrt{1-\epsilon^2}\right)^{\lambda}\sum_{\mu=0}^{n'}\left(a_\mu+b_\mu\right)\rho^{\mu} \tag{2.139}$$

显然，上式在 $r=0$ 时趋于无穷大。

(7) 采用核的点电荷模型时 ($V = -Z/r$)，如果核电荷数 Z 大于光速 (137 a.u.①)，则 Dirac 方程不一定存在束缚态解。假定一个核电荷数为 Z 的单电子原子，由 (2.137) 式可得原子能量为

$$E \approx m_0 c^2 \left[1 - \frac{Z^2\alpha^2}{2n^2} - \frac{Z^4\alpha^4}{2n^3}\left(\frac{1}{|k|} - \frac{3}{4n}\right)\right] \tag{2.140}$$

式中，$\alpha = 1/c$。

由上式可知，一般情况下 $0 < E < m_0 c^2$；当 $Z > c$ 时，上式可能出现 $E < 0$，即非电子束缚态解。解决方法是用核的有限电荷密度分布模型代替点电荷模型 (见 2.4 节)。

2.4　有限核电荷密度分布模型

作为一级近似，原子核可看成是球形，我们常使用“核半径”来表示核的大小。与原子中的电子云不同，原子核有较为确定的表面，即核物质空间分布的几率密度在表面附近从内到外陡然下降为零。许多实验数据表明，原子核的大小基本上和它所含的质子和中子数成正比。

$$R = 1.2A^{1/3}$$

或

$$R = 0.836A^{1/3} + 0.57 \tag{2.141}$$

式中，R 为核半径；A 为原子核的质量数。在相对论量子化学计算中经常使用的是上式中的第二式。总的来说 R 仅在几个费米②左右，远远小于原子半径 (纳米③数量级)，所以在量子化学计算中一般不考虑原子核的大小，将其近似为点电荷，由此得到的氢原子相对论波函数在原点 ($r = 0$) 存在奇点 (见 2.3.3 节)。这一节我们介绍几种球对称有限核电荷密度分布模型，这些模型都是从实验结果中总结出来的。

1. 球面电荷密度分布

由于原子核有较为确定的表面，可以假定核电荷 $\rho(r)$ 均匀分布在半径为 R 的球面上

$$\rho(r) = \begin{cases} \dfrac{Z}{4\pi r^2} & (r = R) \\ 0 & (r \neq R) \end{cases} \tag{2.142}$$

① 1 a.u. = 2.1877×10^6 m·s^{-1}
② 1 费米 (fm) = 10^{-15} 米 (m)
③ 1 纳米 (nm) = 10^{-9} 米 (m)

式中，Z 为核电荷；R 为核半径 (见 (2.141) 式)。与此电荷密度分布对应的静电势能在空间中不连续

$$V(r) = \begin{cases} -\dfrac{Z}{R} & (0 \leqslant r \leqslant R) \\ -\dfrac{Z}{r} & (r > R) \end{cases} \tag{2.143}$$

2. 均匀电荷密度分布

假定核电荷 $\rho(r)$ 在半径为 R' 的球内均匀分布

$$\rho(r) = \begin{cases} \dfrac{3Z}{4\pi R'^3} & (r \leqslant R') \\ 0 & (r > R') \end{cases} \tag{2.144}$$

参数 R' 与核半径 R 的关系为

$$R' = \sqrt{5/3}R \tag{2.145}$$

此时的静电势能在空间中是连续的

$$V(r) = \begin{cases} -\dfrac{3Z}{2R'}\left(1 - \dfrac{r^2}{3R'^2}\right) & (0 \leqslant r \leqslant R') \\ -\dfrac{Z}{r} & (r > R') \end{cases} \tag{2.146}$$

3. Fermi 电荷密度分布

以上两种模型的电荷密度分布在 $r > R$ 时突然变为零，Fermi 电荷密度分布随着 r 的增加逐渐衰减为零

$$\rho(r) = \frac{\rho_0}{1 + \mathrm{e}^{(r-R)/b}} \quad (b > 0) \tag{2.147}$$

式中，ρ_0 为归一化常数；b 为待定参数。对应的势能很难用解析函数来表达，因而在实际计算中只能用数值法求解相应积分。当 $b \to 0$ 时，Fermi 电荷密度分布与均匀电荷密度分布一致。

4. Gauss 电荷密度分布

与 Fermi 电荷密度分布类似，Gauss 电荷密度也是随 r 增加而逐渐衰减至零，但后者是用 Gauss 函数来表达

$$\rho(r) = \rho_0 \mathrm{e}^{-r^2/R^2} \tag{2.148}$$

对应的势能为

$$V(r) = -\frac{Z}{r}\mathrm{erf}(r/R) \tag{2.149}$$

$\mathrm{erf}(x)$ 为 Gauss 误差函数，定义为

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \mathrm{d}t \tag{2.150}$$

以上几个模型中，从计算的角度来看，只有 Gauss 函数的积分最为容易。就计算结果的精度而言，几种模型的差别并不大[5]。此外，还有的电荷密度分布模型采取更复杂的函数形式，感兴趣的读者可阅读 Andrae 的综述文献[6]。

若用有限核电荷密度分布模型取代点电荷模型，计算所得的轨道能量和电子总能量将上升，单电子波函数 (轨道或旋量) 在近核区域的表现也会发生改变，从而影响相对论效应的计算 (相对论效应主要来自近核区域)。对非相对论计算，采用有限核电荷密度分布模型后，径向波函数在原点的尖点将不再存在；对相对论计算而言，有限核电荷密度分布模型可避免径向函数的奇点问题 (见 2.3.3 节)。

参 考 文 献

[1] Greiner W. Theoretical Physics Series 3, Relativistic Quantum Mechanics: Wave Equations. Berlin: Springer-Verlag, 1990.

[2] Greiner W. Theoretical Physics Series 1, Quantum Mechanics: An Introduction. Berlin: Springer-Verlag, 1989.

[3] Greiner W, Reinhardt J. Theoretical Physics Series 4, Quantum Electrodynamics. Berlin: Springer-Verlag, 1992.

[4] Baerends E J, Schwarz W H E, Schwerdtfeger P, et al. Relativistic atomic orbital contraction and expansions: magnitudes and expansions. J Phys B: At Mol Opt Phys, 1990, 23: 3225～3241.

[5] Dyall K G, Knut F Jr. Introduction to Relativistic Quantum Chemistry. New York: Oxford University Press, 2007.

[6] Andrae D. Nuclear charge density distribution in quantum chemistry. // Schwerdtfeger P. Relativistic Electronic Structure Theory, Part 1: Fundamentals, Theoretical and Computational Chemistry. Vol 11. Amsterdam: Elsevier Science B V, 2002: 203～250.

第 3 章　相对论的电磁相互作用

我们知道，分子是构成物质的微小单元，它是能够独立存在并保持物质原有的一切化学性质的最小微粒。分子由原子构成，将这些原子联系起来的就是电磁相互作用。在非相对论级别，即光速无穷大时，两个带电粒子 q_1, q_2 间存在 Coulomb 相互作用

$$V_{12} = \frac{q_1 q_2}{4\pi\epsilon_0 r} \tag{3.1}$$

第 2 章讨论氢原子的 Dirac 方程时，电子和原子核间的相互作用就采用了以上形式 $\left(\frac{-e_1^2}{r}\right)$。相对论的带电粒子间相互作用的严格推导需要用到量子电动力学的知识。为简单计算，我们将在非相对论的电动力学框架内，引入光速的有限性，导出低速条件下 ($u \ll c$) 带电粒子间的相互作用形式，再和量子电动力学的结果进行比较。

本章首先介绍描述电磁场变化规律的 Maxwell 方程组，从而引入电磁场的矢势、标势、规范变换等概念。然后讨论运动电荷产生的电磁场、运动电荷在电磁场中的相对论势能，最后得到运动电荷间相互作用的相对论表达。

3.1　Maxwell 方程组

Maxwell 方程组 (Maxwell equations) 是电磁学中最基本的一个方程组，它反映了一般情况下电荷电流激发电磁场以及电磁场内部运动的规律。对真空中的点电荷，方程可写为

$$\begin{aligned}
\nabla \cdot \boldsymbol{B} &= 0 \\
\nabla \times \boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t} \\
\nabla \cdot \boldsymbol{E} &= \frac{\rho}{\epsilon_0} \\
\nabla \times \boldsymbol{B} &= \frac{1}{c^2}\left[\frac{\boldsymbol{j}}{\epsilon_0} + \frac{\partial \boldsymbol{E}}{\partial t}\right]
\end{aligned} \tag{3.2}$$

方程组中，$\boldsymbol{B}$ 表示磁感应强度；$\boldsymbol{E}$ 表示电场强度；ρ 为电荷密度；ϵ_0 为真空介电常量；$\boldsymbol{j}$ 表示电流密度。方程组中的第一个和第三个方程又称为 Gauss 定理的微分形式，第一式表示磁场的散度为零，说明电流激发的磁场都是无源的。第三式说明电

场的散度和电荷的关系。方程组 (3.2) 式中第二个方程 (Faraday 电磁感应定律) 和第四个方程 (Maxwell 位移电流假设) 则揭示变化电磁场的规律，即变化磁场激发电场 (第二个方程)、变化电场激发磁场 (第四个方程)。

电流密度 $\boldsymbol{j}$ 和电荷密度 ρ 满足连续性方程

$$\nabla \cdot \boldsymbol{j} + \frac{\partial \rho}{\partial t} = 0 \tag{3.3}$$

定义四维矢量 $\boldsymbol{J}$ 和 $\Box$ (读为 quabla)

$$\begin{aligned} \boldsymbol{J} &= (\boldsymbol{j}, \mathrm{i}c\rho) = (j_x, j_y, j_z, \mathrm{i}c\rho) \\ \Box &= \left(\nabla, \frac{-\mathrm{i}}{c}\frac{\partial}{\partial t}\right) \end{aligned} \tag{3.4}$$

$\Box$ 的平方称为 d'Alembert 符号

$$\Box^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \tag{3.5}$$

利用 (3.4) 式，连续性方程 (3.3) 式重写为

$$\Box \cdot \boldsymbol{J} = 0 \tag{3.6}$$

Maxwell 方程组和连续性方程均满足 Einstein 的相对性原理，或称 Lorentz 变换不变性，即方程中每一项按 Lorentz 变换从一个惯性参考系变换到另一个惯性参考系时 (见第 1 章 1.3 节)，方程形式保持不变。

3.2 电磁场的矢势和标势及规范变换

当考虑由电荷分布激发电磁场问题时，引入势的概念来描述电磁场比较方便。由磁场 $\boldsymbol{B}$ 的无源性 (见方程组 (3.2) 式的第一式) 引入矢势 $\boldsymbol{a}$，使

$$\boldsymbol{B} = \nabla \times \boldsymbol{a} \tag{3.7}$$

由附录 B 的 (B.20) 式知 $\nabla \cdot (\nabla \times \boldsymbol{a}) = 0$。矢势 $\boldsymbol{a}$ 的物理意义是：在任一时刻，$\boldsymbol{a}$ 沿任一闭合回路的线积分等于该时刻通过回路内的磁通量。把 (3.7) 式代入 (3.2) 方程组的第二式，得

$$\nabla \times \left(\boldsymbol{E} + \frac{\partial \boldsymbol{a}}{\partial t}\right) = 0 \tag{3.8}$$

该式表示 $\boldsymbol{E} + \dfrac{\partial \boldsymbol{a}}{\partial t}$ 是无旋场，由附录 B.3 中的第三个定理可知，无旋场可表示为标量场 (用 $-\varphi$ 表示) 的梯度，所以有

$$\boldsymbol{E} + \frac{\partial \boldsymbol{a}}{\partial t} = -\nabla\varphi \tag{3.9}$$

一般情况下电场的表示式为

$$\boldsymbol{E}=-\frac{\partial \boldsymbol{a}}{\partial t}-\nabla\varphi \tag{3.10}$$

我们用 (3.7) 式和 (3.9) 式就可以把电磁场用矢势和标势表示出来，显然，这种表示并不唯一 —— 给定的 $\boldsymbol{E},\boldsymbol{B}$ 并不对应于唯一的 $\boldsymbol{a}$ 和 φ。这是因为我们可以将一个任意函数的梯度加到矢势 $\boldsymbol{a}$ 上，得到新的矢势 $\boldsymbol{a}'$，而 $\boldsymbol{a}'$ 和 $\boldsymbol{a}$ 均对应于相同的 $\boldsymbol{B}$，这个加在 $\boldsymbol{a}$ 上的梯度部分在 (3.10) 式中又可以从 $\nabla\varphi$ 中除去，结果也不影响 $\boldsymbol{E}$。

设 ψ 为任意函数，作变换

$$\begin{aligned}\boldsymbol{a}'&=\boldsymbol{a}+\nabla\psi\\ \varphi'&=\varphi-\frac{\partial\psi}{\partial t}\end{aligned} \tag{3.11}$$

有

$$\begin{aligned}\nabla\times\boldsymbol{a}'&=\nabla\times\boldsymbol{a}=\boldsymbol{B}\\ -\frac{\partial\boldsymbol{a}'}{\partial t}-\nabla\varphi'&=-\frac{\partial\boldsymbol{a}}{\partial t}-\nabla\varphi=\boldsymbol{E}\end{aligned} \tag{3.12}$$

即 $(\boldsymbol{a}',\varphi')$ 与 $(\boldsymbol{a},\varphi)$ 描述同一电磁场。变换 (3.11) 式称为势的规范变换。每一组 $(\boldsymbol{a},\varphi)$ 称为一种规范。当势作规范变换时，其他物理量和物理规律保持不变，这种不变性称为规范不变性。从数学上说，规范变换自由度的存在是由于在势的定义式 (3.7) 式和 (3.9) 式中只给出了 $\boldsymbol{a}$ 的旋度，而没有给出 $\boldsymbol{a}$ 的散度①。电磁场 $\boldsymbol{E}$ 和 $\boldsymbol{B}$ 本身对 $\boldsymbol{a}$ 的散度没有任何限制，我们可以取 $\nabla\cdot\boldsymbol{a}$ 为任意的值，每一种选择就对应一种规范。采用适当的辅助条件可以使基本方程和计算简化，而且物理意义也较明显。从计算方便考虑，在不同问题中可以采用不同辅助条件。目前应用最广的是以下两种规范条件：

(1) Lorentz 规范　辅助条件为

$$\nabla\cdot\boldsymbol{a}+\frac{1}{c^2}\frac{\partial\varphi}{\partial t}=0 \tag{3.13}$$

随后我们将看到，采用这种规范时，势的基本方程写为特别简单的对称形式，其物理意义也特别明显。

定义四维矢量 $\boldsymbol{A}$

$$\boldsymbol{A}=\left(\boldsymbol{a},\frac{\mathrm{i}}{c}\varphi\right) \tag{3.14}$$

联合 (3.4) 式，Lorentz 规范简写为

$$\Box\cdot\boldsymbol{A}=0 \tag{3.15}$$

① 旋度和散度的定义见附录 B.3

(2) Coulomb 规范　辅助条件为

$$\nabla \cdot \boldsymbol{a} = 0 \tag{3.16}$$

在 Coulomb 规范中 $\boldsymbol{a}$ 为无源场，因而电场表示式 (3.10) 式中第一项 $-\dfrac{\partial \boldsymbol{a}}{\partial t}$ 是无源场 (即横场)，而第二项 $-\nabla\varphi$ 为无旋场 (即纵场)。这个规范的特点是 $\boldsymbol{E}$ 的纵场完全由 φ 描述，而横场部分由 $\boldsymbol{a}$ 描述，这种划分对讨论某些问题是方便的。显然，当标势 φ 与时间无关时， Coulomb 规范和 Lorentz 规范一致。

现在由 Maxwell 方程组推导矢势 $\boldsymbol{a}$ 和标势 φ 所满足的基本方程。将 (3.7) 式和 (3.10) 式代入 Maxwell 方程组 (见 (3.2) 式) 中的第三、四个方程，得

$$\begin{aligned} \nabla^2\varphi + \frac{\partial}{\partial t}(\nabla \cdot \boldsymbol{a}) &= -\frac{\rho}{\epsilon_0} \\ \nabla^2\boldsymbol{a} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{a}}{\partial t^2} - \nabla\left(\nabla\cdot\boldsymbol{a} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t}\right) &= -\frac{1}{c^2}\frac{\boldsymbol{j}}{\epsilon_0} \end{aligned} \tag{3.17}$$

即为 $\boldsymbol{a}$ 和 φ 所满足的基本方程，适用于一般规范。

若采用 Coulomb 规范，将 (3.16) 式代入上式，得

$$\begin{aligned} \nabla^2\varphi &= -\frac{\rho}{\epsilon_0} \\ \nabla^2\boldsymbol{a} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{a}}{\partial t^2} - \frac{1}{c^2}\frac{\partial}{\partial t}\nabla\varphi &= -\frac{1}{c^2}\frac{\boldsymbol{j}}{\epsilon_0} \end{aligned} \tag{3.18}$$

第一式称为 Poisson 方程。Coulomb 规范的特点是标势满足的方程与静电场情形相同，其解是 Coulomb 势，解出 φ 后代入第二式可解出 $\boldsymbol{a}$。

若采用 Lorentz 规范，将 (3.13) 式代入 (3.17) 式，有

$$\begin{aligned} \nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} &= -\frac{\rho}{\epsilon_0} \\ \nabla^2\boldsymbol{a} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{a}}{\partial t^2} &= -\frac{1}{c^2}\frac{\boldsymbol{j}}{\epsilon_0} \end{aligned} \tag{3.19}$$

用这种规范时 $\boldsymbol{a}$ 和 φ 的方程具有相同形式，其意义也十分明显。(3.19) 式也称为 d′Alembert 方程，它是非齐次的波动方程，其自由项为电流密度和电荷密度。

(3.19) 式告诉我们电荷产生标势波动，电流产生矢势波动。离开电荷电流分布区域后，矢势和标势都以波动形式在空间中传播。采用 d′Alembert 符号 (见 (3.5) 式)，将方程 (3.19) 式重写为

$$\begin{aligned} \Box^2\varphi &= -\frac{\rho}{\epsilon_0} \\ \Box^2\boldsymbol{a} &= -\frac{1}{c^2}\frac{\boldsymbol{j}}{\epsilon_0} \end{aligned} \tag{3.20}$$

利用四维矢量 $\boldsymbol{A}$, $\boldsymbol{J}$ (定义见 (3.14) 式和 (3.4) 式)，d′Alembert 方程可简写为

$$\Box^2 \boldsymbol{A} = -\frac{1}{c^2}\frac{\boldsymbol{J}}{\epsilon_0} \tag{3.21}$$

3.3　运动电荷产生的电磁场

3.2 节我们定义了四维矢量 $\boldsymbol{A} = \left(\boldsymbol{a}, \frac{\mathrm{i}}{c}\varphi\right)$，如果我们在两个不同的惯性参考系 Σ, Σ' 进行测量，分别得到 $\boldsymbol{A}, \boldsymbol{A}'$，二者之间存在什么变换关系呢？

考虑到 Maxwell 方程组满足 Lorentz 变换不变性 (见 3.1 节)，我们自然希望 $\boldsymbol{A}$ 的 Lorentz 变换形式与四维位置矢量 $\boldsymbol{r} = (x, y, z, \mathrm{i}ct)$(见第 1 章 1.2 节) 相同。

设参考系 Σ' 相对于参考系 Σ 以速度 u 运动且两参考系的 x 轴和 x' 轴都沿相对运动方向 (见图 3.1)，参照 (1.13) 式，得不同参考系中测得的 $\boldsymbol{A}, \boldsymbol{A}'$ 的变换关系为

$$\begin{aligned}
a'_x &= \gamma\left(a_x - \frac{u\varphi}{c^2}\right)\\
a'_y &= a_y\\
a'_z &= a_z\\
\varphi' &= \gamma(\varphi - u a_x)
\end{aligned} \tag{3.22}$$

其中，γ 为 Lorentz 因子 (见 (1.14) 式)。

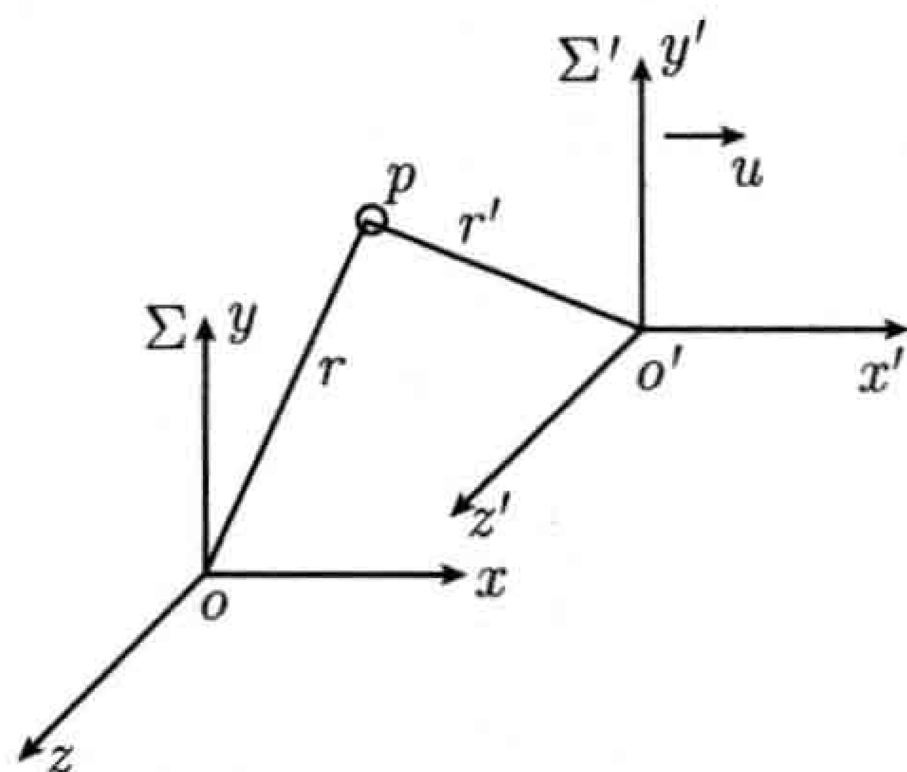


图 3.1　空间点 p 在两个相对运动参考系 Σ, Σ' 中的表示

接下来，我们利用上述条件求得运动点电荷 q 产生的电磁场。假定电荷 q 以速度为 u 沿参考系 Σ 的 x 轴方向运动，另一参考系 Σ' 原点与点电荷重合，x' 轴正方向即为速度的方向 (见图 3.1)。在参考系 Σ' 观察，电荷处于静止状态，空间任一点 p 的电磁场标势即为 Coulomb 势

$$\varphi' = \frac{q}{4\pi\epsilon_0 r'} \tag{3.23}$$

式中，r' 为点 p 到电荷的距离。

在没有外加磁场的条件下，由 Maxwell 方程组 (3.2) 式的第四个方程知该点的矢势 $\boldsymbol{a}'=0$，代入 (3.22) 式，得

$$\begin{aligned} a_x &= \frac{\gamma u \varphi'}{c^2} \\ a_y &= 0 \\ a_z &= 0 \\ \varphi &= \gamma\varphi' \end{aligned} \tag{3.24}$$

联合 (3.23) 式，只要我们知道 r'，便能求出 a_x, φ。

如图 3.1 所示，在参考系 Σ' 观察，由于电磁波传播速度为光速 c，考虑到光速大小的有限性，在时刻 $t'=0$ 时到达 p 点的电磁信号是电荷 q 在较早时刻 $t'=-r'/c$ 激发的。同样，在另一参考系 Σ 观测，时刻 $t=0$ 时点 p 的电磁信号是电荷 q 在较早时刻 $t=-r/c$ 激发的。运用 Lorentz 变换 (见 (1.13) 式) 得

$$t' = \gamma(t - ux/c^2) \tag{3.25}$$

考虑更一般的情形，即 $\boldsymbol{u}$ 为任意方向，由 (1.15) 式知上式中的 ux 替换为 $\boldsymbol{u}\cdot\boldsymbol{r}$。将 $t'=-r'/c,\ t=-r/c$ 代入上式，整理得

$$r' = \gamma\left(r + \frac{\boldsymbol{u}\cdot\boldsymbol{r}}{c}\right) \tag{3.26}$$

代入 (3.23) 式得到 φ'，再代入 (3.24) 式，得到参考系 Σ 中测得的运动电荷 q 的矢势和标势分别为

$$\boldsymbol{a} = \frac{\boldsymbol{u}\varphi}{c^2}, \qquad \varphi = \frac{q}{4\pi\epsilon_0}\left(r + \frac{\boldsymbol{u}\cdot\boldsymbol{r}}{c}\right)^{-1} \tag{3.27}$$

(3.27) 式称为 Liénard-Wiechert 势。注意上式右边各项对应于电磁信号在观测时刻 $t=0$ 之前发出，这说明空间某点 p 的电磁场值不是依赖于同一时刻的电荷电流分布，而是取决于较早时刻的电荷电流分布。反过来说，电荷产生的物理作用不能够立即传到点 p，而是在较晚时刻才传到点 p，所推迟的时间 r/c 正是电磁作用从电荷源点传到 p 所需的时间，c 是电磁作用传播的速度。因此 Liénard-Wiechert 势也称为推迟势。

3.4 运动电子在电磁场中的相对论势能

设点电荷 q 以速度 $\boldsymbol{u}$ 在电磁场中运动，则场对电荷的作用力为

$$\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{u}\times\boldsymbol{B}) \tag{3.28}$$

上式称为 Lorentz 力公式。式中，$\boldsymbol{E}$, $\boldsymbol{B}$ 分别为电场强度和磁感应强度。由电动力学知，电荷 q 在电磁场中的势能为

$$V = q\varphi - q\boldsymbol{u}\cdot\boldsymbol{a} \tag{3.29}$$

$\boldsymbol{a}$, φ 分别为该电磁场的矢势和标势。它们和电场强度 $\boldsymbol{E}$、磁感应强度 $\boldsymbol{B}$ 的关系见 (3.7) 式和 (3.9) 式。

Lorentz 力 $\boldsymbol{F}$ 和 V 的关系可写为

$$\boldsymbol{F} = -\nabla V + \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V}{\partial u_i}\hat{\boldsymbol{e}}_i \quad (i = x, y, z) \tag{3.30}$$

式中，$\hat{\boldsymbol{e}}_i(i = x, y, z)$ 为沿 x, y, z 轴的单位矢量。

有兴趣的读者可自已证明，由 (3.30) 式定义的力 $\boldsymbol{F}$ 满足 Lorentz 力公式 (3.28) 式。

3.5　运动电荷间相互作用的相对论形式

3.4 节讨论了运动电荷在电磁场中的势能，假定该电磁场是由电荷 q_2(速度为 $\boldsymbol{u}_2$) 产生的，将 (3.27) 式代入 (3.29) 式，得电荷 q_1(速度为 $\boldsymbol{u}_1$) 在电磁场中的势能为

$$V_{12} = \frac{q_1 q_2}{4\pi\epsilon_0}\left(r + \frac{\boldsymbol{u}_2\cdot\boldsymbol{r}}{c}\right)^{-1}\cdot\left(1 - \frac{\boldsymbol{u}_1\cdot\boldsymbol{u}_2}{c^2}\right) \tag{3.31}$$

V_{12} 称为两个电荷的相互作用能。

反过来，我们也可以认为电荷 q_2(速度为 $\boldsymbol{u}_2$) 在由电荷 q_1(速度为 $\boldsymbol{u}_1$) 产生的电磁场中运动，同样可得两个电荷的相互作用能

$$V_{21} = \frac{q_1 q_2}{4\pi\epsilon_0}\left(r + \frac{\boldsymbol{u}_1\cdot\boldsymbol{r}}{c}\right)^{-1}\cdot\left(1 - \frac{\boldsymbol{u}_1\cdot\boldsymbol{u}_2}{c^2}\right) \tag{3.32}$$

由于相互作用能表达式对于下标 1, 2 互换并不具有对称性，一般情况下 $V_{12} \neq V_{21}$ ——与两个运动电荷的相互作用能是唯一的相矛盾。为此，我们需要寻找电荷相互作用的对称表达形式。

假定参考系 Σ' 相对于参考系 Σ 以速度 $\boldsymbol{u}$ 运动，将 (1.15) 式重新整理，得位置矢量的 Lorentz 变换为

$$\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{u}\left[(\gamma - 1)\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{u^2} - \gamma t\right] \tag{3.33}$$

式中，γ 为 Lorentz 因子 (见 (1.14) 式)。

已知在 $t_0 = 0$ 时刻观测到的空间某点的电磁场信号是在 $t = -r/c$ 时刻由电荷发出的 (见 3.3 节)，考虑低速的情形，即 $u \ll c$，近似认为 $t \approx 0$，代入 (3.33) 式，得

$$\boldsymbol{r}' \approx \boldsymbol{r} + \boldsymbol{u}\left[(\gamma - 1)\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{u^2}\right] \tag{3.34}$$

由此可以算出 $\boldsymbol{r}'$ 的模

$$\begin{aligned}|\boldsymbol{r}'|^2 &= \boldsymbol{r}'\cdot\boldsymbol{r}' \\ &\approx r^2\left[1 + \left(\frac{\gamma\boldsymbol{u}\cdot\boldsymbol{r}}{cr}\right)^2\right]\end{aligned} \tag{3.35}$$

代入 (3.24) 式，得

$$\varphi = \gamma\varphi' = \frac{\gamma q}{4\pi\epsilon_0 r'} \approx \frac{\gamma q}{4\pi\epsilon_0 r}\left[1 + \left(\frac{\gamma\boldsymbol{u}\cdot\boldsymbol{r}}{cr}\right)^2\right]^{-1/2} \tag{3.36}$$

考虑到 $u \ll c$，将 Lorentz 因子作级数展开

$$\gamma^2 = \frac{1}{1 - \dfrac{u^2}{c^2}} = 1 + \frac{u^2}{c^2} + \cdots \tag{3.37}$$

忽略展开式中小于 u^2/c^2 的高阶展开项，只保留前两项，得 $\gamma^2 \approx 1 + \dfrac{u^2}{c^2}$，代入 (3.36) 式并略去小项 ($< u^2/c^2$)，有

$$\varphi \approx \frac{q}{4\pi\epsilon_0 r}\left[1 + \left(\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{cr}\right)^2 - \frac{u^2}{c^2}\right]^{-1/2} \tag{3.38}$$

将上式右边作 Taylor 级数展开，得

$$\varphi \approx \frac{q}{4\pi\epsilon_0 r}\left[1 + \frac{u^2}{2c^2} - \frac{1}{2}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{cr}\right)^2 + O\left(\frac{u^4}{c^4}\right)\right] \tag{3.39}$$

$O\left(\dfrac{u^4}{c^4}\right)$ 表示此项大小约在 $\dfrac{u^4}{c^4}$ 数量级。将 (3.39) 式代入 (3.27) 式，得矢势表达式

$$\boldsymbol{a} = \frac{\boldsymbol{u}\varphi}{c^2} \approx \frac{q}{4\pi\epsilon_0 rc}\left[\frac{u}{c} + O\left(\frac{u^3}{c^3}\right)\right] \tag{3.40}$$

将此式与 (3.39) 式一起代入 (3.29) 式，得 q_1，q_2 的相互作用能为

$$V_{12} = \frac{q_1 q_2}{4\pi\epsilon_0 r}\left[1 + \frac{u_2^2}{2c^2} - \frac{1}{2}\left(\frac{\boldsymbol{u}_2\cdot\boldsymbol{r}}{cr}\right)^2 + O\left(\frac{u_2^4}{c^4}\right)\right] - \frac{q_1 q_2}{4\pi\epsilon_0 rc^2}\boldsymbol{u}_1\cdot\boldsymbol{u}_2 - O\left(\frac{u_1 u_2^3}{c^4}\right) \tag{3.41}$$

上式关于下标 1, 2 的互换依然不对称。

在 3.2 节提到，Coulomb 规范下得到 Coulomb 势 (见 (3.18) 式)，由此得到的电荷–电荷相互作用能关于电荷指标互换是对称的。注意到代入 (3.29) 式的运动电荷的电磁场 (见 (3.27) 式) 是在 Lorentz 规范下得到的 (见 3.3 节)，为得到关于电荷指标互换对称的电荷–电荷相互作用能表达式，我们需要寻找 Coulomb 规范下运动电荷的电磁场。

引入规范函数 Λ，由 (3.11) 式，得

$$\begin{aligned}\boldsymbol{a}^{\mathrm{C}} &= \boldsymbol{a}^{\mathrm{L}} + \nabla\Lambda \\ \varphi^{\mathrm{C}} &= \varphi^{\mathrm{L}} - \frac{\partial\Lambda}{\partial t}\end{aligned} \tag{3.42}$$

上标 C, L 分别代表 Coulomb 规范和 Lorentz 规范。已知 Coulomb 规范下 $\varphi^{\mathrm{C}} = \dfrac{q}{4\pi\epsilon_0 r}$，令 $\Lambda = \dfrac{q}{4\pi\epsilon_0}\lambda$，代入 (3.42) 式，并考虑到 (3.39) 式，得

$$\frac{\partial\lambda}{\partial t} = \frac{1}{2c^2}\left[\frac{u^2}{r} - \frac{(\boldsymbol{u}\cdot\boldsymbol{r})^2}{r^3}\right] \tag{3.43}$$

上式中我们已略去 u^2/c^2 的高阶小项。将式中的 u^2 以 $\boldsymbol{u}\cdot\dfrac{\mathrm{d}(-\boldsymbol{r})}{\mathrm{d}t}$ 代替。注意 $\boldsymbol{r}$ 的方向由电荷指向空间某点 p，$\mathrm{d}(-\boldsymbol{r})/\mathrm{d}t$ 代表电荷的速度 $\boldsymbol{u}$。忽略电荷运动的加速度，即假定 $\mathrm{d}\boldsymbol{u}/\mathrm{d}t \approx 0$，(3.43) 式可进一步改写为

$$\frac{\partial\lambda}{\partial t} = -\frac{1}{2c^2}\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{r}\right] \tag{3.44}$$

由上式容易求出 λ

$$\lambda = -\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{2c^2 r} = -\frac{\boldsymbol{u}\cdot\boldsymbol{n}}{2c^2} \tag{3.45}$$

式中，$\boldsymbol{n}$ 表示沿 $\boldsymbol{r}$ 的单位矢量。将 (3.45) 式代入 (3.42) 式，得

$$\boldsymbol{a}^{\mathrm{C}} = \boldsymbol{a}^{\mathrm{L}} - \frac{q}{8\pi\epsilon_0 c^2}\nabla(\boldsymbol{u}\cdot\boldsymbol{n}) \tag{3.46}$$

而

$$\begin{aligned}\nabla(\boldsymbol{u}\cdot\boldsymbol{n}) &= \frac{1}{r}\left[\boldsymbol{u} - \boldsymbol{n}(\boldsymbol{u}\cdot\boldsymbol{n})\right] \\ &= \frac{\boldsymbol{u}}{r} - \frac{\boldsymbol{r}(\boldsymbol{u}\cdot\boldsymbol{r})}{r^3}\end{aligned} \tag{3.47}$$

上式的推导用到了 (B.11) 式。将 (3.47) 式代入 (3.46) 式，并略去小于 u^2/c^2 的各项，得

$$\boldsymbol{a}^{\mathrm{C}} = \frac{q}{8\pi\epsilon_0 rc}\left[\frac{\boldsymbol{u}}{c} + \frac{\boldsymbol{r}(\boldsymbol{u}\cdot\boldsymbol{r})}{cr^2}\right] \tag{3.48}$$

将 Coulomb 规范下得到的 φ^{C} 和 $\boldsymbol{a}^{\mathrm{C}}$ 代入 (3.29) 式，得到关于电荷下标互换对称的 q_1, q_2 的相互作用能表达式

$$V_{12} = \frac{q_1 q_2}{4\pi\epsilon_0 r}\left[1 - \frac{\boldsymbol{u}_1\cdot\boldsymbol{u}_2}{2c^2} - \frac{(\boldsymbol{u}_1\cdot\boldsymbol{r})(\boldsymbol{u}_2\cdot\boldsymbol{r})}{2c^2r^2}\right] \tag{3.49}$$

至此，在经典电动力学的框架内，我们推出了相对论的电荷–电荷相互作用能表达式。式中第一项与非相对论条件下的电荷–电荷 Coulomb 相互作用一致，第二项和第三项就像是来自电荷之间电流–电流、电流–偶极子、偶极子–偶极子的相互作用。在推导过程中我们假定电荷运动速度远小于光速，省略了 u^2/c^2 的高阶小项。

考虑一个多电子的孤立体系 (如原子或分子)，体系内存在着核–核、电子–核及电子–电子三种电荷相互作用，在核冻结近似下 (即忽略核的运动)，由 (3.49) 式知核–核及电子–核相互作用仍为经典 Coulomb 相互作用，相对论贡献为零。

根据量子电动力学，电子–电子相互作用是通过光子交换进行的，大小和光子频率有关， Coulomb 规范下电子–电子相互作用的严格形式为[1]

$$\hat{g}_\omega(i,j) = \frac{1}{r_{ij}} - \frac{\hat{\boldsymbol{\alpha}}_i\cdot\hat{\boldsymbol{\alpha}}_j}{r_{ij}}\mathrm{e}^{\mathrm{i}r_{ij}\omega_{ij}/c} + \frac{(\hat{\boldsymbol{\alpha}}_i\cdot\nabla_i)(\hat{\boldsymbol{\alpha}}_j\cdot\nabla_j)}{\omega_{ij}^2 r_{ij}c^2}\left(\mathrm{e}^{\mathrm{i}r_{ij}\omega_{ij}/c} - 1\right) \tag{3.50}$$

交换光子的频率等于轨道能级的差值，即 $\omega_{ij} = |\epsilon_i - \epsilon_j|$ 。在低能 ($\epsilon \ll c^2$) 条件下将上式中的指数函数 $\mathrm{e}^{\mathrm{i}r_{ij}\omega_{ij}/c}$ 作级数展开，略去小于 $1/c^2$ 的高阶展开项，得电子–电子 Coulomb-Breit(CB) 相互作用

$$\hat{g}_{\mathrm{CB}}(i,j) = \frac{1}{r_{ij}}\left[1 - \frac{\hat{\boldsymbol{\alpha}}_i\cdot\hat{\boldsymbol{\alpha}}_j}{2} - \frac{(\hat{\boldsymbol{\alpha}}_i\cdot\hat{\boldsymbol{r}}_{ij})(\hat{\boldsymbol{\alpha}}_j\cdot\hat{\boldsymbol{r}}_{ij})}{2r_{ij}^2}\right] \tag{3.51}$$

式中，第一项为 Coulomb 相互作用；第二、三项合称为 Breit 相互作用；第二项又称为 Gaunt (磁) 相互作用。由于 Breit 相互作用由微扰理论导出，在实际应用中适用于作微扰处理[2]，但是，如果将变分过程限制在电子态 (见第 5 章 5.1 节)，Breit 相互作用也可以用于变分计算[3]，研究表明两种处理方法得到的结果差别并不大。

另外，我们可用量子电动力学的速度表达式 $c\hat{\boldsymbol{\alpha}}$ 代替经典速度 $\boldsymbol{u}$，代入 (3.49) 式并取原子单位① 得

$$\hat{V}_{12} = \frac{1}{r}\left[1 - \frac{\hat{\boldsymbol{\alpha}}_1\cdot\hat{\boldsymbol{\alpha}}_2}{2} - \frac{(\hat{\boldsymbol{\alpha}}_1\cdot\hat{\boldsymbol{r}})(\hat{\boldsymbol{\alpha}}_2\cdot\hat{\boldsymbol{r}})}{2r^2}\right] \tag{3.52}$$

以上结果和量子电动力学的结果 (3.51) 式完全一致。

将 (3.52) 式代入第 2 章 (2.44) 式，得多电子体系的 Dirac-Coulomb-Breit(DCB) Hamilton 算符

$$\hat{H} = \sum_i^n \hat{h}_{\mathrm{D}}(i) + \sum_{i<j}^n \hat{g}(i,j) + \sum_{\lambda<\mu}^N \frac{Z_\lambda Z_\mu}{r_{\lambda\mu}} \tag{3.53}$$

① 原子单位的定义见附录C

式中，各项均取原子单位；n 代表体系电子数目；N 代表原子核数目；$\hat{h}_{\mathrm{D}}(i)$ 为

$$\hat{h}_{\mathrm{D}}(i) = c\hat{\boldsymbol{\alpha}} \cdot \hat{\boldsymbol{p}}_i + \hat{\beta}c^2 - \sum_{\lambda} \frac{Z_\lambda}{r_i} \tag{3.54}$$

$\hat{g}(i,j)$ 即为 (3.51) 式。DCB Hamilton 算符是目前用于分子计算的最精确 Hamilton 算符 (精确到 α^2 数量级，α 为精细结构常数，约为 1/137)。如果 (3.53) 式中 $\hat{g}(i,j)$ 取 Coulomb 相互作用，则得到 Dirac-Coulomb(DC) Hamilton 算符；如果取 Coulomb 和 Gaunt 相互作用两项 (即 (3.51) 式中的前两项)，则得到 Dirac-Coulomb-Gaunt(DCG) Hamilton 算符。从量子电动力学 (quantum electrodynamics，QED) 的理论出发，还能得到更高阶相对论校正如真空极化、电子的自能等 (见附录 D)，通常称为 QED 效应。对量子电动力学有兴趣的读者可参阅 Greiner 和 Reinhardt 编写的《量子电动力学》一书[2]。关于 QED 对元素化学性质的影响读者可参阅 Pyykkö 的综述文章 [4]。

参 考 文 献

[1] Berestetskii V B, Lifshitz E M, Pitaevskii L P. Relativistic Quantum Theory. Oxford: Pergamon Press, 1971.

[2] Greiner W, Reinhardt J. Theoretical Physics Series 4, Quantum Electrodynamics. Berlin: Springer-Verlag, 1992.

[3] Indelicato P, Santos J P, Boucard S, et al. QED and relativistic corrections in superheavy elements. Eur Phys J D, 2007, 45: 155~170.

[4] Pyykkö P. The physics behind chemistry and the periodic table. Chem Rev, 2012, 112: 379~394.

第 4 章　相对论的对称性原理

对称性原理是物理学和化学中的一个核心概念，它决定了分子的结构、光谱、晶体结构等。体系的很多性质也只与体系的对称性质有关，与体系的其他特殊性无关。所以了解一个体系的对称性会有助于我们预测它的理化性质。此外，当要对体系的性质进行定量计算时，利用对称性原理可以使计算工作量大为减少。本章目的是介绍处理相对论量子化学的对称性时需要用到的理论工具及探讨 Dirac 方程的对称性。

在数学上，对称性由群论来表述。群论是代数学的一个分支，已有近两百年的历史。现在群论在物理学和化学中已成为一种不可缺少的数学工具。代数学上群论研究的是有限或无限集合中定义的特定代数运算的性质，包含的内容十分广泛，群论在物理和化学中的运用只是其中很小的一部分。本章只介绍群论中与相对论量子化学有关的部分，而且认为读者熟悉有限群及其表示论的基本知识。

我们将首先介绍由三维位形空间的位置对称变换而导致的态函数 (或称量子态) 及算符变换的基础知识，再由总角动量为半整数的量子态在绕任一轴的 360° 旋转变换下改变符号这一性质引入二倍群和相对论的分子对称性，接着讨论 Dirac 方程的对称不变性，重点是描述自旋空间和位形空间转动的两个无限连续 Lie 群，即二维特殊酉群 SU(2) 和三维特殊实正交群 SO(3)。

4.1　空间对称变换

位置变换是在三维位形空间即我们所在的物理空间从一个点到另一个点的变换。变换 $\hat{Q}$ 是一个三维位形空间中的算符，它将点 $\boldsymbol{r}$ 变为另一个点 $\boldsymbol{r}'$，记为

$$\boldsymbol{r}' = \hat{Q}\boldsymbol{r} \tag{4.1}$$

若每一个 $\boldsymbol{r}$, $\boldsymbol{r}'$ 都有确定的值，则变换 $\hat{Q}$ 就有了明确的定义。我们所说的变换 $\hat{Q}$，是不改变任何两点距离的那些变换。对一个原子或分子体系，若位置变换的一个集合 $\{\hat{Q}_i\}$ $(i = 1, 2, 3, \cdots)$ 是保持该体系不变的对称变换，则这个集合必构成一个群，称为这个系统的对称变换群。$\hat{Q}_i$ 也称为该群的元素，且满足

(1) 封闭性，任意两个元素的乘积都是集合中的一个元素。这里群的乘法规定为相继的变换，$\hat{Q}_1\hat{Q}_2$ 是先进行变换 $\hat{Q}_2$，接着进行变换 $\hat{Q}_1$。

(2) 乘法满足结合律，如 $(\hat{Q}_1\hat{Q}_2)\hat{Q}_3 = \hat{Q}_1(\hat{Q}_2\hat{Q}_3)$。

(3) 对于任一元素 $\hat{Q}_i$，必然存在它的逆元素，记为 $\hat{Q}_i^{-1}$。由于我们已经预先设定变换 $\hat{Q}$ 不改变任何两点距离，这就保证了逆元素的存在。

以上所列实际上就是群的定义。对称变换群的阶可以是有限的，也可以是无穷的。

1. 态函数的变换

考虑一个由单粒子组成的系统，其态函数在位置表象中的表示为 $\phi(\boldsymbol{r}) = \langle \boldsymbol{r}|\phi\rangle$。现在用位置变换 $\hat{Q}$ 把这一函数值的分布整体移到另一个地方，也就是说当 $\boldsymbol{r}$ 受到 $\hat{Q}$ 的作用变到 $\boldsymbol{r}' = \hat{Q}\boldsymbol{r}$ 点时，带着它的函数值一直到 $\boldsymbol{r}'$ 去。这时系统的态发生了改变，成为 $|\phi'\rangle$，而其位置表象的态函数成为 $\phi'(\boldsymbol{r}) = \langle \boldsymbol{r}|\phi\rangle'$。显然，新函数在新点的值等于老函数在老点上的值，即态函数变换的条件是

$$\phi'(\hat{Q}\boldsymbol{r}) = \phi(\boldsymbol{r}) \tag{4.2}$$

由此知

$$\phi'(\boldsymbol{r}) = \phi(\hat{Q}^{-1}\boldsymbol{r}) \tag{4.3}$$

这就是位置变换 $\hat{Q}$ 导致的态函数的变换。新老态函数的关系可以用一个函数空间的算符 $\hat{R}(\hat{Q})$ 来表示

$$\phi'(\boldsymbol{r}) = \hat{R}(\hat{Q})\phi(\boldsymbol{r}) = \phi(\hat{Q}^{-1}\boldsymbol{r}) \tag{4.4}$$

由于 $\hat{Q}$ 不改变任意两点的距离，$\phi'(\boldsymbol{r})$ 和 $\phi(\boldsymbol{r})$ 二者只是地点和方位不同，而尺度和形状不变，因而不影响其归一化，于是可得 $\hat{R}^\dagger(\hat{Q})\hat{R}(\hat{Q}) = 1$，所以 $\hat{R}(\hat{Q})$ 为一酉算符。

2. 算符的变换

位置变换 $\hat{Q}$ 既导致了态函数的变换，也导致了算符的变换。假定一般算符 $\hat{w}$，变换后所得新算符记为 $\hat{w}'$，有

$$\hat{R}(\hat{w}\phi) = \hat{w}'\phi' = \hat{w}'\hat{R}\phi \tag{4.5}$$

上式左乘 $\hat{R}^{-1}$，得

$$\hat{w}\phi = \hat{R}^{-1}\hat{w}'\hat{R}\phi \tag{4.6}$$

比较等式两边，得新算符 $\hat{w}'$ 与老算符 $\hat{w}$ 的关系为

$$\hat{w}' = \hat{R}\hat{w}\hat{R}^{-1} \tag{4.7}$$

现在讨论位置算符 $\hat{\boldsymbol{r}}$，用变换算符 $\hat{R}(\hat{Q})$ 作用，得新位置算符 $\hat{\boldsymbol{r}}' = \hat{R}\hat{\boldsymbol{r}}\hat{R}^{-1}$，将该算符作用到任一态函数 $\phi(\boldsymbol{r})$ 并应用 (4.4) 式，有

$$\hat{\boldsymbol{r}}'\phi(\boldsymbol{r}) = \hat{R}\hat{\boldsymbol{r}}\hat{R}^{-1}\phi(\boldsymbol{r}) = \hat{R}\hat{\boldsymbol{r}}\phi(\hat{Q}\boldsymbol{r}) = \hat{Q}^{-1}\hat{\boldsymbol{r}}\phi(\hat{Q}^{-1}\hat{Q}\boldsymbol{r}) = \hat{Q}^{-1}\hat{\boldsymbol{r}}\phi(\boldsymbol{r}) \tag{4.8}$$

由上式有

$$\hat{\boldsymbol{r}}' = \hat{R}\hat{\boldsymbol{r}}\hat{R}^{-1} = \hat{Q}^{-1}\hat{\boldsymbol{r}} \tag{4.9}$$

此式是位置算符 $\hat{\boldsymbol{r}}$ 的变换关系。注意此式与 (4.1) 式的区别，(4.1) 式是位形空间中位置矢量的变换关系，而 (4.9) 式是位置算符的变换关系。在 (4.9) 式中 $\hat{\boldsymbol{r}}$ 兼有算符和位置矢量的两种身份

$$\hat{\boldsymbol{r}} = \sum_{i=1}^{3} r_i \hat{\boldsymbol{e}}_i \tag{4.10}$$

(4.9) 式中 $\hat{R}$ 只对系数 r_i 作用，$\hat{Q}^{-1}$ 只对单位矢量 $\hat{\boldsymbol{e}}_i$ 发生作用。

4.2 相对论的单电子原子的对称性

4.2.1 量子态的转动

先考虑非相对论情况，即无自旋粒子绕 z 轴的转动。设体系绕 z 轴转过一个无限小角度 $\delta\phi$，粒子角坐标从 $\phi \to \phi' = \phi + \delta\phi$ ，波函数从 $\Psi \to \Psi' = \hat{R}_z(\delta\phi)\Psi$，试求 $\hat{R}_z(\delta\phi)$。利用函数的级数展开，可将 Ψ' 写为

$$\begin{aligned}\Psi'(\phi) &= \Psi(\phi + \delta\phi) \\ &= \Psi(\phi) + \delta\phi\frac{\partial}{\partial\phi}\Psi + \frac{1}{2!}(\delta\phi)^2\frac{\partial^2}{\partial\phi^2}\Psi + \cdots \\ &= \mathrm{e}^{-\mathrm{i}\delta\phi\hat{l}_z/\hbar}\Psi\end{aligned} \tag{4.11}$$

式中，$\hat{l}_z = -\mathrm{i}\hbar\partial/\partial\phi$，是轨道角动量算符 $\hat{\boldsymbol{l}}$ 的 z 分量。这样，无自旋粒子绕 z 轴旋转 $\delta\phi$ 角的算符可表示成

$$\hat{R}_z(\delta\phi) = \mathrm{e}^{-\mathrm{i}\delta\phi\hat{l}_z/\hbar} \tag{4.12}$$

设体系绕 z 轴旋转一个有限角 ϕ，它可以看成体系相继进行一系列无限小角度旋转的总的效果，因而

$$\hat{R}_z(\phi) = \mathrm{e}^{-\mathrm{i}\phi\hat{l}_z/\hbar} \tag{4.13}$$

在相对论级别，轨道角动量应代替为总角动量 $\hat{\boldsymbol{J}}$ ，即

$$\hat{R}_z(\phi) = \mathrm{e}^{-\mathrm{i}\phi\hat{J}_z/\hbar} \tag{4.14}$$

更普遍地讲，粒子绕空间任一方向 $\boldsymbol{n}$ 旋转一个角度 ϕ 时量子态的转动算符为

$$\hat{R}_{\boldsymbol{n}}(\phi) = \mathrm{e}^{-\mathrm{i}\phi\boldsymbol{n}\cdot\hat{\boldsymbol{J}}/\hbar} \tag{4.15}$$

4.2.2 单电子原子的 360° 转动

在第 2 章我们求解了氢原子的 Dirac 方程时，找到了描述氢原子每一个量子态的四个量子数 (n, k, j, m_j)，氢原子的总角动量本征函数可用 Dirac 符号表示为

$|j, m_j\rangle$。假定氢原子绕 z 轴转动有限角 ϕ，由于径向函数与转动无关，只需考虑总角动量波函数绕 z 轴的转动，代入 (4.14) 式得转动后的新函数为

$$\hat{R}_z(\phi)|j, m_j\rangle = \mathrm{e}^{-\mathrm{i}\phi m_j}|j, m_j\rangle \tag{4.16}$$

当转动角为 2π 时，由于 m_j 始终为半整数，所以有

$$\hat{R}_z(2\pi)|j, m_j\rangle = -|j, m_j\rangle \tag{4.17}$$

在非相对论级别，氢原子角动量本征函数用 $|l, m_l\rangle$ 表示，将其绕 z 轴转动 360°，考虑到 m_l 始终为整数，得

$$\hat{R}_z(2\pi)|l, m_l\rangle = \mathrm{e}^{-\mathrm{i}2\pi m_l}|l, m_l\rangle = |l, m_l\rangle \tag{4.18}$$

比较 (4.17) 式和 (4.18) 式知，考虑相对论效应时，360° 旋转操作将改变氢原子 (或单电子原子) 波函数的符号。更进一步推广，所有自旋为半整数体系的波函数在旋转 360° 操作下将改变符号，这一新的对称操作用 $\bar{E}$ 表示。

4.3 分子对称群：二倍群

在非相对论级别，分子的对称性用有限点群来描述，其中并不包括对称操作 $\bar{E}$。为包含这一新的对称操作 $\bar{E}$，相对论的分子对称群定义 G' 为非相对论点群 G 和群 $\{E, \bar{E}\}$ 的直积

$$G' = G \otimes \{E, \bar{E}\} \tag{4.19}$$

这意味着新群 G' 的对称操作数目将是非相对论点群 G 的两倍，群的类和不可约表示的数目也将增加，我们称 G' 为二倍群。二倍群 G' 的特征标表可以从非相对论点群 G 的特征标表按以下规则推出。以下所列规则是群表示理论中关于不可约表示及特征标表的几条基本规则，证明可参见 Cotton 编著的 *Chemical Applications of Group Theory*[1]。

(1) 所有不可约表示的维数 (l_i) 的平方和等于群的阶 (h)，即群中对称操作的数目。也就是说

$$\sum_{i=1}^{n} l_i^2 = l_1^2 + l_2^2 + \cdots = h \tag{4.20}$$

式中，n 表示不可约表示的数目。

(2) 任一不可约表示的所有特征标 $\chi_i(R)$ 的平方和等于群的阶，即

$$\sum_R \chi_i^2(R) = h \tag{4.21}$$

式中，下标 i 代表不可约表示；R 代表对称操作。

(3) 由不同不可约表示的特征标组成的矢量是互相正交的。即

$$\sum_R \chi_i(R)\chi_j(R) = 0 \quad (i \neq j) \tag{4.22}$$

(4) 属于同类的所有对称操作的特征标相同。

(5) 同一群中，不可约表示的数目等于类的数目。

下面我们以相对论二倍群 C_1' 和 C_{2v}' 为例，说明如何从非相对论点群 C_1, C_{2v} 出发，得到 C_1', C_{2v}' 的特征标表，更多二倍群的特征标表可参见文献 [2, 3]。

非相对论点群 C_1 仅含恒等操作 $\{E\}$(见表 4.1)，由 (4.19) 式得二倍群 C_1' 的对称操作有 $\{E,\ \bar{E}\}$，为得到 C_1' 的特征标表 (见表 4.1)，我们首先将对称操作进行分类。根据群论中关于类的定义[1]，E, $\bar{E}$ 属于不同的两个类，参考以上规则的第 (5) 条和第 (1) 条，得 C_1' 有两个一维不可约表示 (分别用 A, B 表示)。A 为全对称不可约表示，此时两个对称操作 E, $\bar{E}$ 的特征标都为 1；对不可约表示 B，由于恒等操作 E 的特征标在任何表示下均为 1，利用 (4.22) 式，得 B 表示下 $\bar{E}$ 的特征标为 -1，从 4.2.2 节知自旋为 1/2 的粒子在 $\bar{E}$ 操作下改变符号，所以该不可约表示可记为 $B_{1/2}$(见表 4.1)。

表 4.1　C_1 群 (左) 和 C_1' 群 (右) 的特征标表

C_1	E
A	1

C_1'	E	$\bar{E}$
A	1	1
$B_{1/2}$	1	-1

再来看非相对论点群 C_{2v}，该群有四个对称操作：恒等操作 E、一个二重轴 C_2 和两个通过二重轴的对称面 $\sigma_v(xz)$, $\sigma_v'(yz)$(见表 4.2)。这四个对称操作与 $\{E,\ \bar{E}\}$ 的直积生成二倍群 C_{2v}'，此群包含八个对称操作 $\{E,\ \bar{E},\ C_2,\ C_2\bar{E},\ \sigma_v(xz),\ \sigma_v(xz)\bar{E},\ \sigma_v'(yz),\ \sigma_v'(yz)\bar{E}\}$。

表 4.2　C_{2v} 群的特征标表

C_{2v}	E	C_2	$\sigma_v(xz)$	$\sigma_v'(yz)$
A_1	1	1	1	1
A_2	1	1	-1	-1
B_1	1	-1	1	-1
B_2	1	-1	-1	1

由以上规则第 (1) 条 (见 (4.20) 式) 知，当且仅当 n 的取值为 5，且 $l_1 = l_2 = l_3 = l_4 = 1$, $l_5 = 2$ 时，$\sum_{i=1}^{5} l_i^2 = 8$。也就是说二倍群 C_{2v}' 包含 4 个一维和 1 个二维

共 5 个不可约表示。

对称操作 E, C_2, $\sigma_v(xz)$, $\sigma'_v(yz)$ 在 4 个一维不可约表示下的特征标可直接从非相对论的点群 C_{2v} 得到。恒等操作 E 在二维不可约表示下的特征标为 2，利用以上规则第 (3) 条可得 C_2, $\sigma_v(xz)$, $\sigma'_v(yz)$ 在二维不可约表示下的特征标均为 0。

接着考虑 360° 旋转操作 $\bar{E}$，此操作可分解为两次 180° 旋转操作，即 $\bar{E}=C_2C_2$，参考表 4.2 中二重轴 C_2 的特征标，得 $\bar{E}$ 在 4 个一维不可约表示下的特征标均为 1，再利用规则第 (3) 条和恒等操作 E 的特征标，可得 $\bar{E}$ 在二维不可约表示下的特征标为 -2，与其他 4 类对称操作的特征标相比得对称操作 $\bar{E}$ 自成一类。一般来说，二倍转动群中 360° 旋转操作 $\bar{E}$ 与恒等操作 E 一样，均自成一类。由规则第 (5) 条知群中类的数目和不可约表示的数目相等，均为 5。已有 E, C_2, $\sigma_v(xz)$, $\sigma'_v(yz)$, $\bar{E}$ 五个操作各成一类，剩下未归类的三个对称操作 $C_2\bar{E}$, $\sigma_v(xz)\bar{E}$, $\sigma'_v(yz)\bar{E}$ 则必和以上某种操作同类。

将这三个操作进行如下分解

$$
\begin{aligned}
C_2\bar{E} &= C_2C_2C_2 \\
\sigma_v(xz)\bar{E} &= C_2\sigma_v(xz)C_2 \\
\sigma'_v(yz)\bar{E} &= C_2\sigma'_v(yz)C_2
\end{aligned} \tag{4.23}
$$

由类的定义[1] 知 $C_2\bar{E}$ 和 C_2, $\sigma_v(xz)\bar{E}$ 和 $\sigma_v(xz)$, $\sigma'_v(yz)\bar{E}$ 和 $\sigma'_v(yz)$ 分别属同类对称操作，它们的特征标也相同。由此我们得到了 C'_{2v} 群的特征标表 (见表 4.3)。新增加的二维不可约表示用符号 $E_{1/2}$ 表示，1/2 表示自旋。

表 4.3 C'_{2v} 群的特征标表

C'_{2v}	E	C_2	$\sigma_v(xz)$	$\sigma'_v(yz)$	$\bar{E}$
		$C_2\bar{E}$	$\sigma_v(xz)\bar{E}$	$\sigma'_v(yz)\bar{E}$	
A_1	1	1	1	1	1
A_2	1	1	−1	−1	1
B_1	1	−1	1	−1	1
B_2	1	−1	−1	1	1
$E_{1/2}$	2	0	0	0	−2

4.4 位形空间中的转动和 SO(3) 群

分子或原子的结构是用三维位形空间 ($\mathbb{R}_3$) 来描述的，该空间中的转动是使所有的位置矢量 $\boldsymbol{r}$ 都绕一过原点的固定轴 $\boldsymbol{n}$ 转过一定角度的变换 (见图 (4.1))。从数

学上看，位形空间中无限小角度转动的表示式为

$$\begin{aligned} \boldsymbol{r}' &= \hat{R}_{\boldsymbol{n}}(\mathrm{d}\phi)\boldsymbol{r} \\ &= \boldsymbol{r} + \mathrm{d}\phi\boldsymbol{n} \times \boldsymbol{r} \end{aligned} \tag{4.24}$$

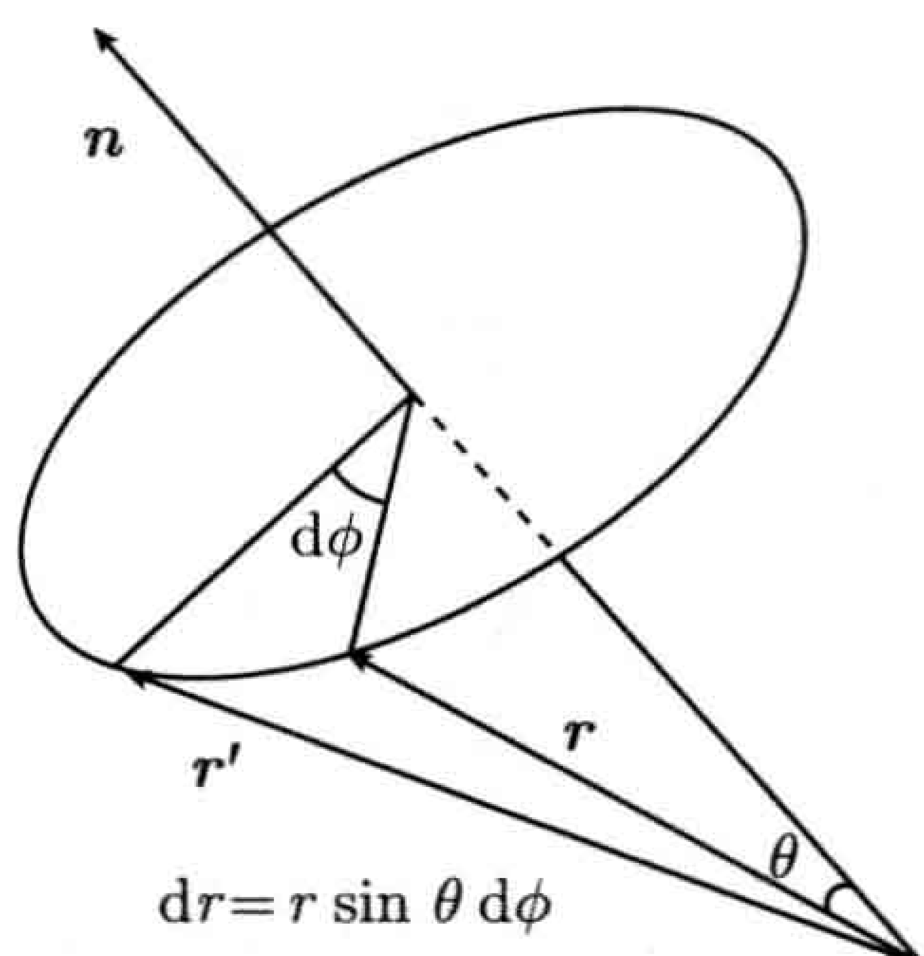


图 4.1 矢量 $\boldsymbol{r}$ 绕固定轴 $\boldsymbol{n}$ 的无限小角度转动

这一点可以从图 4.1 看出。这样的转动不会破坏分子的结构，也就是说

$$\begin{aligned} \langle \hat{R}\boldsymbol{r}|\hat{R}\boldsymbol{r}\rangle &= \langle \boldsymbol{r}|\hat{R}^{\mathrm{T}}\hat{R}|\boldsymbol{r}\rangle = \langle \boldsymbol{r}|\boldsymbol{r}\rangle \\ \langle \hat{R}\boldsymbol{r}|\hat{R}\boldsymbol{r}'\rangle &= \langle \boldsymbol{r}|\hat{R}^{\mathrm{T}}\hat{R}|\boldsymbol{r}'\rangle = \langle \boldsymbol{r}|\boldsymbol{r}'\rangle \end{aligned} \tag{4.25}$$

$\hat{R}^{\mathrm{T}}$ 表示 $\hat{R}$ 的转置，符合以上条件的 $\hat{R}$ 必为 $\mathbb{R}_3$ 空间的正交矩阵，即

$$\hat{R}^{\mathrm{T}}\hat{R} = \boldsymbol{I}_3 \tag{4.26}$$

$\boldsymbol{I}_3$ 表示 3×3 单位矩阵。所有满足 (4.26) 式的 3×3 矩阵的集合构成三维实正交群，用 O(3) 表示。接下来求变换算符 $\hat{R}$。

令 $\boldsymbol{r} = (x, y, z)$ 绕 z 轴转过无限小角度 $\mathrm{d}\phi$，由 (4.24) 式，有

$$\boldsymbol{r}' = \hat{R}_z(\mathrm{d}\phi)\boldsymbol{r} = (x - y\mathrm{d}\phi)\hat{e}_x + (y + x\mathrm{d}\phi)\hat{e}_y + z\hat{e}_z \tag{4.27}$$

式中，$\hat{e}_x, \hat{e}_y, \hat{e}_z$ 分别表示 x, y, z 轴的单位矢量。

由上式可得

$$\hat{R}_z(\mathrm{d}\phi) = \begin{pmatrix} 1 & -\mathrm{d}\phi & 0 \\ \mathrm{d}\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.28}$$

引入矩阵 $-\mathrm{i}X_z$

$$-\mathrm{i}X_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.29}$$

将 (4.28) 式重写为

$$\hat{R}_z(\mathrm{d}\phi) = \boldsymbol{I}_3 - \mathrm{i}\mathrm{d}\phi X_z \tag{4.30}$$

绕 z 轴旋转一个有限角 ϕ 的转动可以看成体系相继进行一系列无限小角度转动的总的效果

$$\hat{R}_z(\phi) = \lim_{n\to\infty}\left[\hat{R}_z\left(\frac{\phi}{n}\right)\right]^n = \lim_{n\to\infty}\left(\boldsymbol{I}_3 - \mathrm{i}\frac{\phi}{n}X_z\right)^n = \mathrm{e}^{-\mathrm{i}X_z\phi} \tag{4.31}$$

我们称 X_z 为旋转生成元。由于 X_z 为矩阵，得 $\mathrm{e}^{-\mathrm{i}X_z\phi}$ 为矩阵指数函数。矩阵指数函数的运算遵循以下基本规则

(1)
$$\mathrm{e}^A = \sum_{k=0}^{\infty}\frac{1}{k!}A^k \tag{4.32}$$

(2)
$$\mathrm{e}^A\mathrm{e}^B = \mathrm{e}^{A+B} \quad ([A,B]=0) \tag{4.33}$$

(3)
$$|\mathrm{e}^A| = \mathrm{e}^{\mathrm{Tr}(A)} \tag{4.34}$$

$\mathrm{Tr}(A)$ 等于矩阵 A 的对角元之和，称为 A 的迹。

(4)
$$B\mathrm{e}^AB^{-1} = \mathrm{e}^{BAB^{-1}} \quad (\text{B 为酉矩阵}) \tag{4.35}$$

对于绕 x 轴或 y 轴的转动，旋转生成元 X_x 或 X_y 的定义与 X_z 类似

$$-\mathrm{i}X_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad -\mathrm{i}X_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \tag{4.36}$$

引入矢量 $\boldsymbol{X}$，其三个分量即为旋转生成元 X_x, X_y, X_z。我们发现，$\boldsymbol{X}$ 的三个分量满足循环对易规则

$$[X_i, X_j] = \mathrm{i}\epsilon_{ijk}X_k \tag{4.37}$$

ϵ_{ijk} 为 Levi-Civita 符号

$$\epsilon_{ijk} = \begin{cases} 1 & (i,j,k) = \{1,2,3\},\{2,3,1\},\{3,1,2\}\ (\text{偶置换}) \\ -1 & (i,j,k) = \{3,2,1\},\{2,1,3\},\{1,3,2\}\ (\text{奇置换}) \\ 0 & i=j\ \text{或}\ j=k\ \text{或}\ k=i \end{cases} \tag{4.38}$$

式中，1, 2, 3 分别代表 x, y, z。对空间中任一矢量 $\boldsymbol{r}$ 绕任一轴 $\boldsymbol{n} = (n_x, n_y, n_z)$ 的有限角转动，我们都可以将它分解成绕 x, y, z 轴的三个基本旋转的复合

$$\boldsymbol{r}' = \hat{R}_{\boldsymbol{n}}(\phi)\boldsymbol{r} = \mathrm{e}^{-\mathrm{i}\boldsymbol{X}\cdot\boldsymbol{n}\phi}\boldsymbol{r} \tag{4.39}$$

由于矩阵 X_x, X_y, X_z 的迹为零，易得 $\hat{R_{\boldsymbol{n}}}(\phi)$ 的行列式为 1

$$|\hat{R}_{\boldsymbol{n}}(\phi)| = \mathrm{e}^{\mathrm{Tr}(-\mathrm{i}\boldsymbol{X}\cdot\boldsymbol{n}\phi)} = 1 \tag{4.40}$$

由 $\hat{R}_{\boldsymbol{n}}(\phi)$ 生成的群是三维实正交群 O(3) 的一个子群，称为三维特殊实正交群，用 SO(3) 表示。

1. **旋转变换 $\hat{R}_{\boldsymbol{n}}(\phi)$ 用三角函数表达**

取 $\hat{R}_{\boldsymbol{n}}(\phi)$(见 (4.39) 式) 的级数展开

$$\mathrm{e}^{-\mathrm{i}\boldsymbol{X}\cdot\boldsymbol{n}\phi} = \sum_{k=0}^{\infty}\frac{1}{k!}\left[-\mathrm{i}\boldsymbol{X}\cdot\boldsymbol{n}\phi\right]^k \tag{4.41}$$

因为

$$\begin{aligned}
\boldsymbol{X}\cdot\boldsymbol{n} &= \mathrm{i}\begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ n_y & n_x & 0 \end{pmatrix} \\
(\boldsymbol{X}\cdot\boldsymbol{n})^2 &= -\begin{pmatrix} n_x^2 & n_xn_y & n_zn_x \\ n_xn_y & n_y^2 & n_zn_y \\ n_zn_x & n_zn_y & n_z^2 \end{pmatrix} \\
&= \boldsymbol{I}_3 - \boldsymbol{n}\boldsymbol{n}^T \\
(\boldsymbol{X}\cdot\boldsymbol{n})^3 &= \boldsymbol{X}\cdot\boldsymbol{n}
\end{aligned} \tag{4.42}$$

所以

$$(\boldsymbol{X}\cdot\boldsymbol{n})^k = \begin{cases} \boldsymbol{X}\cdot\boldsymbol{n} & (k=2m+1, \quad m=0,1,2,\cdots) \\ (\boldsymbol{X}\cdot\boldsymbol{n})^2 & (k=2m, \quad m=1,2,\cdots) \end{cases} \tag{4.43}$$

将以上结果代入 (4.41) 式，并利用三角函数 $\sin\phi$ 和 $\cos\phi$ 的级数展开

$$\begin{aligned}
\sin\phi &= \phi - \frac{1}{6}\phi^3 + \cdots \\
\cos\phi &= 1 - \frac{1}{2}\phi^2 + \cdots
\end{aligned} \tag{4.44}$$

得

$$\begin{aligned}
\mathrm{e}^{-\mathrm{i}\boldsymbol{X}\cdot\boldsymbol{n}\phi} &= \boldsymbol{I}_3 - \boldsymbol{n}\boldsymbol{n}^T + \boldsymbol{n}\boldsymbol{n}^T + \sum_{k=2,4,\cdots}^{\infty}\frac{(\boldsymbol{X}\cdot\boldsymbol{n})^2}{k!}[\mathrm{i}\phi]^k + \sum_{k=1,3,\cdots}^{\infty}\frac{-(\boldsymbol{X}\cdot\boldsymbol{n})}{k!}[\mathrm{i}\phi]^k \\
&= (\boldsymbol{X}\cdot\boldsymbol{n})^2\cos\phi + \mathrm{i}(\boldsymbol{X}\cdot\boldsymbol{n})\sin\phi + \boldsymbol{n}\boldsymbol{n}^T \\
&= \boldsymbol{I}_3\cos\phi + (1-\cos\phi)\boldsymbol{n}\boldsymbol{n}^T + \mathrm{i}(\boldsymbol{X}\cdot\boldsymbol{n})\sin\phi
\end{aligned} \tag{4.45}$$

$\boldsymbol{n}$ 为三维列向量，3×3 矩阵 $\boldsymbol{n}\boldsymbol{n}^T$ 的矩阵元为

$$(\boldsymbol{n}\boldsymbol{n}^T)_{ij} = n_i n_j \tag{4.46}$$

2. 位置算符、动量算符、角动量算符的旋转变换

将上面得到的 $\hat{R}_{\boldsymbol{n}}(\phi)$(见 (4.39) 式) 代入位置算符变换关系 (见 (4.9) 式)，得变换后的新位置算符 $\hat{\boldsymbol{r}}'$ 为

$$\hat{\boldsymbol{r}}' = \hat{Q}^{-1}\hat{\boldsymbol{r}} = \hat{R}_{\boldsymbol{n}}^{-1}(\phi)\hat{\boldsymbol{r}}$$

对动量算符 $\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla$ 和角动量算符 $\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}}\times\hat{\boldsymbol{p}}$，这两个都是位置矢量的算符, 它们与位置算符 $\hat{\boldsymbol{r}}$ 有相同的变换形式

$$\begin{aligned}\hat{\boldsymbol{p}}' &= \hat{R}_{\boldsymbol{n}}^{-1}(\phi)\hat{\boldsymbol{p}}\\ \hat{\boldsymbol{L}}' &= \hat{R}_{\boldsymbol{n}}^{-1}(\phi)\hat{\boldsymbol{L}}\end{aligned} \tag{4.47}$$

至于其他算符，由于坐标 (x,y,z) 通常以复杂的函数形式隐含在算符内，一般不能得到简单的变换形式。

4.5 自旋空间中的转动和 SU(2) 群

4.5.1 自旋态和自旋算符的表示

电子的自旋在某给定方向 (习惯上取为 s_z) 只能取 $\pm\dfrac{\hbar}{2}$ 两个分立值，自旋态的一组正交完备基矢通常用 α 和 β 表示

$$\alpha = \left|\frac{1}{2},\frac{1}{2}\right\rangle = \begin{pmatrix}1\\0\end{pmatrix},\quad \beta = \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \begin{pmatrix}0\\1\end{pmatrix} \tag{4.48}$$

描述自旋态的波函数的一般形式为

$$\eta = a\left|\frac{1}{2},\frac{1}{2}\right\rangle + b\left|\frac{1}{2},-\frac{1}{2}\right\rangle = \begin{pmatrix}a\\b\end{pmatrix} \tag{4.49}$$

$|a|^2$ 表示 $\hat{s}_z = \hbar/2$ 的概率，$|b|^2$ 表示 $\hat{s}_z = -\hbar/2$ 的概率。

电子的自旋角动量算符用 Pauli 算符 $\hat{\boldsymbol{\sigma}}$ 表示，即 $\hat{\boldsymbol{S}} = \dfrac{\hbar}{2}\hat{\boldsymbol{\sigma}}$，有时也称 Pauli 算符为自旋角动量算符。 Pauli 算符的矩阵表示形式见第 2 章 (2.27) 式，其三个分量满足循环对易规则

$$[\hat{\sigma}_i,\hat{\sigma}_j] = 2\mathrm{i}\epsilon_{ijk}\hat{\sigma}_k \tag{4.50}$$

ϵ_{ijk} 为 Levi-Civita 符号 (见 (4.38) 式)。与轨道角动量算符 $\hat{\boldsymbol{L}}$ 类似，自旋角动量 $\hat{\boldsymbol{S}}$ 也存在梯度算符

$$\hat{S}_+ = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix},\quad \hat{S}_- = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix} \tag{4.51}$$

将 $\hat{S}_+, \hat{S}_-$ 分别作用到 α 和 β，得

$$\begin{aligned}\hat{S}_+\left|\frac{1}{2},\frac{1}{2}\right\rangle &= \begin{pmatrix}0\\0\end{pmatrix}\\ \hat{S}_+\left|\frac{1}{2},-\frac{1}{2}\right\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle\\ \hat{S}_-\left|\frac{1}{2},\frac{1}{2}\right\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle\\ \hat{S}_-\left|\frac{1}{2},-\frac{1}{2}\right\rangle &= \begin{pmatrix}0\\0\end{pmatrix}\end{aligned} \tag{4.52}$$

自旋空间的任一自旋算符 $\hat{w}(\boldsymbol{s})$ 可写为

$$\hat{w}(\boldsymbol{s}) = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a} = \frac{\hbar}{2}(\hat{\sigma}_x a_x + \hat{\sigma}_y a_y + \hat{\sigma}_z a_z) \tag{4.53}$$

式中，$\boldsymbol{a}$ 为实矢量。

4.5.2 自旋空间中的转动

考虑自旋空间中的对称变换，用算符 $\hat{\Omega}$ 表示。由对称变换的一般定义知 $\hat{\Omega}$ 不能改变自旋空间中任意两个向量的内积，即

$$\hat{\Omega}\langle\eta|\xi\rangle = \langle\eta|\hat{\Omega}^\dagger\hat{\Omega}|\xi\rangle = \langle\eta|\xi\rangle \tag{4.54}$$

所以必有 $\hat{\Omega}^\dagger\hat{\Omega} = 1$，即 $\hat{\Omega}$ 为酉算符，用 $\hat{U}$ 表示。满足酉算符条件的 $\hat{U}$ 可写为

$$\hat{U} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{w}(\boldsymbol{s})} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}} \tag{4.55}$$

上式的第二个等号利用了 (4.53) 式。容易证明所有 $\hat{U}$ 的集合满足群的定义[1]，称为二维酉群。

由于 Pauli 矩阵 $\hat{\boldsymbol{\sigma}}$ 的迹为零，得算符 $\hat{U}$ 的行列式

$$|\hat{U}| = \mathrm{e}^{\mathrm{Tr}(-\frac{\mathrm{i}}{2}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a})} = 1 \tag{4.56}$$

我们又称所有 $\hat{U}$ 生成的群为二维特殊酉群，用 SU(2) 表示。

自旋空间中绕任一轴 $\boldsymbol{n}$ 转动 ϕ 角的酉变换可写为

$$\hat{U}_{\boldsymbol{n}}(\phi) = \mathrm{e}^{-\frac{\mathrm{i}}{2}\phi\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}} \tag{4.57}$$

将矩阵指数函数 $\hat{U}_{\boldsymbol{n}}(\phi)$ 作级数展开，并利用

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{n})^2 &= (\boldsymbol{n}\cdot\boldsymbol{n})\boldsymbol{I}_2 + \mathrm{i}\boldsymbol{\sigma}\cdot\boldsymbol{n}\times\boldsymbol{n}\\ &= \boldsymbol{I}_2\end{aligned} \tag{4.58}$$

和 (4.44) 式，有

$$\begin{aligned}\hat{U}_{\boldsymbol{n}}(\phi) &= \mathrm{e}^{-\frac{\mathrm{i}}{2}\phi\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\left[-\frac{\mathrm{i}}{2}\phi\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}\right]^k \\ &= \sum_{k=0,2,4,\cdots}^{\infty}\frac{\boldsymbol{I}_2}{k!}\left[\frac{\mathrm{i}}{2}\phi\right]^k + \sum_{k=1,3,5,\cdots}^{\infty}\frac{\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}}{k!}\left[\frac{-\mathrm{i}}{2}\phi\right]^k \\ &= \boldsymbol{I}_2\cos\frac{\phi}{2} - \mathrm{i}(\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}})\sin\frac{\phi}{2}\end{aligned} \tag{4.59}$$

式中，$\boldsymbol{I}_2$ 为 2×2 单位矩阵。

1. Pauli 算符的转动变换

考虑自旋空间内绕 z 轴的转动，将 $\hat{U}_z(\phi)$ 作用到 $\hat{\boldsymbol{\sigma}}$，由 (4.7) 式得新算符 $\hat{\boldsymbol{\sigma}}'$

$$\hat{\boldsymbol{\sigma}}' = \hat{U}_z(\phi)\hat{\boldsymbol{\sigma}}\hat{U}_z^{-1}(\phi) = \mathrm{e}^{-\frac{\mathrm{i}}{2}\phi\sigma_z}\hat{\boldsymbol{\sigma}}\mathrm{e}^{\frac{\mathrm{i}}{2}\phi\sigma_z} \tag{4.60}$$

将 (4.59) 式代入上式，并利用 Pauli 算符分量的循环对易关系 (见 (4.50) 式)，得

$$\hat{\boldsymbol{\sigma}}' = \hat{\boldsymbol{\sigma}}\begin{pmatrix}\cos\phi & -\sin\phi & 0\\ \sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix} = \hat{\boldsymbol{\sigma}}\hat{R}_z(\phi) \tag{4.61}$$

式中，$\hat{\boldsymbol{\sigma}}$ 取行矢量形式，即 $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$。

以上结论可推广到算符 $\hat{\boldsymbol{\sigma}}$ 绕空间任一轴 $\boldsymbol{n}$ 的转动，新算符 $\hat{\boldsymbol{\sigma}}'$ 可写为

$$\hat{\boldsymbol{\sigma}}' = \hat{\boldsymbol{\sigma}}\hat{R}_{\boldsymbol{n}}(\phi) \tag{4.62}$$

2. SO(3) 与 SU(2) 的同态关系

由 (4.53) 式，我们可建立三维位形空间中的位置矢量 $\boldsymbol{r}(x,y,z)$ 与自旋空间中的算符 $\hat{w}$ 的一一对应关系

$$\hat{w} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{r} = \frac{\hbar}{2}\begin{pmatrix} z & x-\mathrm{i}y\\ x+\mathrm{i}y & -z\end{pmatrix} \tag{4.63}$$

将属于 SU(2) 群的任一变换 $\hat{U}$ 作用到算符 $\hat{w}$，得新算符 $\hat{w}'$

$$\hat{w}' = \hat{U}\hat{w}\hat{U}^{-1} \tag{4.64}$$

此时的 $\hat{w}'$ 必对应于位形空间的另一位置向量 $\boldsymbol{r}'$

$$\hat{w}' = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{r}' = \frac{\hbar}{2}\begin{pmatrix} z' & x'-\mathrm{i}y'\\ x'+\mathrm{i}y' & -z'\end{pmatrix} \tag{4.65}$$

由于

$$|\hat{w}'| = |\hat{U}\hat{w}\hat{U}^{-1}| = |\hat{w}| \tag{4.66}$$

所以

$$x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2 \tag{4.67}$$

由此可以看出，每一个 SU(2) 群的元素 $\hat{U}$ 可以把 $\boldsymbol{r}$ 变到 $\boldsymbol{r'}$ 而长度不变。即 $\hat{U}$ 与 SO(3) 中的某一元素等价，记为 $\hat{R}(\hat{U})$。

$$\boldsymbol{r'} = \hat{R}(\hat{U})\boldsymbol{r} \tag{4.68}$$

代入 (4.64) 式，并考虑到 (4.63) 式，得

$$\hat{U}\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{r}\hat{U}^{-1} = \hat{\boldsymbol{\sigma}} \cdot \hat{R}(\hat{U})\boldsymbol{r} \tag{4.69}$$

如果将两个 SU(2) 群的元素的乘积 $\hat{U}_1\hat{U}_2$ 作用到算符 $\hat{w}$，有

$$\begin{aligned}
\left(\hat{U}_1\hat{U}_2\right)\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{r}\left(\hat{U}_1\hat{U}_2\right)^{-1} &= \hat{U}_1\left(\hat{U}_2\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{r}\hat{U}_2^{-1}\right)\hat{U}_1^{-1} \\
&= \hat{U}_1\hat{\boldsymbol{\sigma}} \cdot \hat{R}\left(\hat{U}_2\right)\boldsymbol{r}\hat{U}_1^{-1} \\
&= \hat{\boldsymbol{\sigma}} \cdot \hat{R}\left(\hat{U}_1\right)\hat{R}\left(\hat{U}_2\right)\boldsymbol{r} \\
&= \hat{\boldsymbol{\sigma}} \cdot \hat{R}\left(\hat{U}_1\hat{U}_2\right)\boldsymbol{r}
\end{aligned} \tag{4.70}$$

即 $\hat{U}_1\hat{U}_2$ 与 SO(3) 中的元素 $\hat{R}(\hat{U}_1\hat{U}_2)$ 等价。由此，我们就在 SU(2) 和 SO(3) 之间建立了同态关系。从 (4.64) 式可知，$\hat{U}$ 和 $-\hat{U}$ 与相同的 $\hat{R}$ 对应，也就是说它们之间的对应关系是二对一的关系。例如，在 SU(2) 群中 $\hat{U}_{\boldsymbol{n}}(0) = -\hat{U}_{\boldsymbol{n}}(2\pi)$(见 (4.57) 式)，而在 SO(3) 中 $\hat{R}_{\boldsymbol{n}}(0) = \hat{R}_{\boldsymbol{n}}(2\pi)$。

4.6 Dirac 方程的转动变换

本节讨论自旋和位形空间中的转动 (SU(2) ⊗ SO(3)) 对 Dirac 方程的影响。考虑单电子的 Dirac 定态方程

$$\left(\hat{h}_{\mathrm{D}} + V\right)\Psi = \left[c\hat{\boldsymbol{\alpha}} \cdot \hat{\boldsymbol{p}} + \hat{\beta}m_0c^2 + V\right]\Psi = E\Psi \tag{4.71}$$

这里假定 V 是球对称的静态标势。自旋和位形空间中的转动对称操作可写为 $\hat{U}_{\boldsymbol{n_1}}(\phi_1)\hat{R}_{\boldsymbol{n_2}}(\phi_2)$，将其作用到 $\hat{h}_{\mathrm{D}} + V$ ，由 (4.7) 式得新算符

$$\begin{aligned}
(\hat{h}_{\mathrm{D}} + V)' &= \hat{U}_{\boldsymbol{n_1}}(\phi_1)\hat{R}_{\boldsymbol{n_2}}(\phi_2)\left(\hat{h}_{\mathrm{D}} + V\right)\hat{R}_{\boldsymbol{n_2}}^{-1}(\phi_2)\hat{U}_{\boldsymbol{n_1}}^{-1}(\phi_1) \\
&= c\hat{\boldsymbol{\alpha}}' \cdot \hat{\boldsymbol{p}}' + \hat{\beta}' + V'
\end{aligned} \tag{4.72}$$

根据我们的假定，转动对称操作不改变球对称的 V，即 $V'=V$。由于 $\hat{\beta}$ 为对角常数矩阵 (见 (2.36) 式)，易得 $\hat{\beta}'=\hat{\beta}$。再来看 $\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}$

$$\hat{\boldsymbol{\alpha}}'\cdot\hat{\boldsymbol{p}}'=\left[\hat{U}_{\boldsymbol{n_1}}(\phi_1)\hat{\boldsymbol{\alpha}}\hat{U}^{-1}_{\boldsymbol{n_1}}(\phi_1)\right]\cdot\left[\hat{R}_{\boldsymbol{n_2}}(\phi_2)\hat{\boldsymbol{p}}\hat{R}^{-1}_{\boldsymbol{n_2}}(\phi_2)\right] \tag{4.73}$$

由 (2.36) 式、(4.61) 式和 (4.47) 式，得

$$\begin{aligned}\hat{\boldsymbol{\alpha}}'&=\hat{\boldsymbol{\alpha}}\hat{R}_{\boldsymbol{n_1}}(\phi_1)\\ \hat{\boldsymbol{p}}'&=\hat{R}^{-1}_{\boldsymbol{n_2}}(\phi_2)\hat{\boldsymbol{p}}\end{aligned} \tag{4.74}$$

代入 (4.73) 式，得

$$\hat{\boldsymbol{\alpha}}'\cdot\hat{\boldsymbol{p}}'=\hat{\boldsymbol{\alpha}}\hat{R}_{\boldsymbol{n_1}}(\phi_1)\hat{R}^{-1}_{\boldsymbol{n_2}}(\phi_2)\hat{\boldsymbol{p}}=\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}\quad(\boldsymbol{n_1}=\boldsymbol{n_2},\quad\phi_1=\phi_2) \tag{4.75}$$

由此我们得到结论，自旋和位形空间的转动变换 $\hat{U}_{\boldsymbol{n}}(\phi)\hat{R}_{\boldsymbol{n}}(\phi)$ 保持 $\hat{h}_{\mathrm{D}}+V$ 不变。

接下来讨论多电子体系。以分子为例，分子的 Dirac Hamilton 算符可写为

$$\hat{H}=\sum_{i}^{n}\hat{h}_{\mathrm{D}}(i)+\sum_{i<j}^{n}\hat{g}(i,j)+\sum_{\lambda<\mu}^{N}\frac{Z_\lambda Z_\mu}{4\pi\epsilon_0 r_{\lambda\mu}} \tag{4.76}$$

式中各项意义见 (3.53) 式。对电子–电子相互作用项 $\hat{g}(i,j)$

$$\hat{g}(i,j)=\frac{1}{r_{ij}}\left[1-\frac{\hat{\boldsymbol{\alpha}}_i\cdot\hat{\boldsymbol{\alpha}}_j}{2}-\frac{(\hat{\boldsymbol{\alpha}}_i\cdot\hat{\boldsymbol{r}}_{ij})(\hat{\boldsymbol{\alpha}}_j\cdot\hat{\boldsymbol{r}}_{ij})}{2r_{ij}^2}\right] \tag{4.77}$$

显然，上式中的 Coulomb 相互作用项在转动对称操作下保持不变。Breit 相互作用项包含 $\hat{\boldsymbol{\alpha}}_i\cdot\hat{\boldsymbol{\alpha}}_j$ 和 $\hat{\boldsymbol{\alpha}}_{i(j)}\cdot\hat{\boldsymbol{r}}$ 两项，二者在 $\mathrm{SU}(2)\otimes\mathrm{SO}(3)$ 变换下均保持不变。对核–核相互作用 (见 (4.76) 式中最后一项)，为使该项在转动对称操作下保持不变，我们可以将位形空间的转动操作 $\hat{R}_{\boldsymbol{n}}(\phi)$ 限制为保持分子空间结构对称性的转动轴，这些对称轴构成 SO(3) 群的一个子群，用 G 表示。Dirac 算符在属于 $(\mathrm{SU}(2)\otimes\mathrm{G})$ 群的对称操作 $\hat{U}_{\boldsymbol{n}}(\phi)\hat{R}_{\boldsymbol{n}}(\phi)$ 下保持不变。

4.7 空间反演

在三维位形空间中，空间反演变换的定义是改变所有空间坐标的符号，空间反演算符可写为

$$\hat{I}\boldsymbol{r}=-\boldsymbol{I}_3\boldsymbol{r}=-\boldsymbol{r} \tag{4.78}$$

$\boldsymbol{I}_3$ 为 3×3 单位矩阵。显然 $-\boldsymbol{I}_3$ 也属于三维实正交群 O(3)，但由于 $|-\boldsymbol{I}_3|=-1$，所以 $-\boldsymbol{I}_3$ 并不属于 SO(3) 群。

函数空间中的空间反演算符 $\hat{I}$ 定义为改变函数中所有坐标的符号

$$\hat{I}\Psi(\boldsymbol{r}) = \Psi(-\boldsymbol{r}) \tag{4.79}$$

空间反演算符又称宇称算符，在函数空间的表达形式为

$$\hat{I}\Psi(\boldsymbol{r}) = \Psi(-\boldsymbol{r}) = \begin{cases} \Psi(\boldsymbol{r}) & (\text{偶宇称}) \\ -\Psi(\boldsymbol{r}) & (\text{奇宇称}) \\ \text{其他情况} & (\text{无确切宇称}) \end{cases} \tag{4.80}$$

我们将相对论空间反演算符写为自旋和位形空间两个变换算符的积

$$\hat{I} = \hat{a}\hat{I}_{\mathrm{R}} \tag{4.81}$$

式中，$\hat{I}_{\mathrm{R}}$ 为位形空间的反演算符；$\hat{a}$ 待定。将 $\hat{I}$ 作用到自由电子的 Dirac Hamilton 算符，得

$$\hat{h}'_{\mathrm{D}} = \hat{I}\hat{h}_{\mathrm{D}}\hat{I}^{-1} = c\left(\hat{a}\hat{\boldsymbol{\alpha}}\hat{a}^{-1}\right)\cdot\left(\hat{I}_{\mathrm{R}}\hat{\boldsymbol{p}}\hat{I}_{\mathrm{R}}^{-1}\right) + mc^2\left(\hat{a}\hat{\beta}\hat{a}^{-1}\right) \tag{4.82}$$

由于 $\hat{I}_{\mathrm{R}}\hat{\boldsymbol{p}}\hat{I}_{\mathrm{R}}^{-1} = -\hat{\boldsymbol{p}}$，代入上式得

$$\hat{h}'_{\mathrm{D}} = -c\left(\hat{a}\hat{\boldsymbol{\alpha}}\hat{a}^{-1}\right)\cdot\hat{\boldsymbol{p}} + mc^2\left(\hat{a}\hat{\beta}\hat{a}^{-1}\right) \tag{4.83}$$

假定 Dirac Hamilton 算符在空间反演下不变，即 $\hat{h}'_{\mathrm{D}} = \hat{h}_{\mathrm{D}}$，则有

$$\begin{aligned} -\hat{\boldsymbol{\alpha}} &= \hat{a}\hat{\boldsymbol{\alpha}}\hat{a}^{-1} \\ \hat{\beta} &= \hat{a}\hat{\beta}\hat{a}^{-1} \end{aligned} \tag{4.84}$$

不难发现，当 $\hat{a} = \hat{\beta}$ 时上式成立。由此确定相对论反演算符

$$\hat{I} = \hat{\beta}\hat{I}_{\mathrm{R}} \tag{4.85}$$

将 $\hat{I}$ 作用到二分量波函数，得

$$\hat{I}\begin{pmatrix} \Psi_{\mathrm{L}} \\ \Psi_{\mathrm{S}} \end{pmatrix} = \begin{pmatrix} \hat{I}_{\mathrm{R}} & 0 \\ 0 & -\hat{I}_{\mathrm{R}} \end{pmatrix}\begin{pmatrix} \Psi_{\mathrm{L}} \\ \Psi_{\mathrm{S}} \end{pmatrix} = \begin{pmatrix} \hat{I}_{\mathrm{R}}\Psi_{\mathrm{L}} \\ -\hat{I}_{\mathrm{R}}\Psi_{\mathrm{S}} \end{pmatrix} \tag{4.86}$$

以氢原子为例，氢原子波函数可写为 (见第 2 章)

$$\Psi_{jmk} = \begin{pmatrix} \Psi_{\mathrm{L}} \\ \Psi_{\mathrm{S}} \end{pmatrix} = \begin{pmatrix} \dfrac{F(r)}{r}\chi_{jm}^{1,2} \\ \dfrac{-\mathrm{i}G(r)}{r}\chi_{jm}^{2,1} \end{pmatrix} \tag{4.87}$$

Ψ_{L} 和 Ψ_{S} 的宇称由 $\chi^{1,2}$ 的宇称决定，即由球谐函数 $Y_{l_1}^m$ 和 $Y_{l_2}^m$ 决定 (见 (2.101) 式)，球谐函数的宇称则由角量子数 l 决定

$$\hat{I}_{\mathrm{R}} Y_l^m = (-1)^l Y_l^m \tag{4.88}$$

所以 Ψ_{L} 和 Ψ_{S} 的宇称由 l_1, l_2 的奇偶性决定。由 (2.97) 式和 (2.98) 式可知，无论量子数 k 小于零还是大于零，l_1, l_2 的奇偶性正好相反，即若 l_1 为奇则 l_2 为偶，反之亦然。所以 Ψ_{L} 和 Ψ_{S} 有相反的宇称

$$\hat{I}_{\mathrm{R}} \Psi_{\mathrm{L}} = \pm 1, \qquad \hat{I}_{\mathrm{R}} \Psi_{\mathrm{S}} = \mp 1 \tag{4.89}$$

整个波函数的宇称和大分量波函数 Ψ_{L} 相同，所以通常用 Ψ_{L} 的对称性来表示四分量波函数的对称性。

4.8 反映和旋转反映

1. 对称面

如果有一个通过物体的平面，物体对这一平面的镜像与其自身重合，则说这个平面是物体的一个对称面。在非相对论的有限点群理论中，对这一平面的反映操作 σ_q 写为

$$\sigma_q = \hat{I}_{\mathrm{R}} C_{2q} \tag{4.90}$$

式中，平面 q 决定垂直于该平面的一个二重轴，如对称面是 xy 平面，则二重轴为 z 轴。与此类似，二倍群中的反映操作定义为相对论的空间反演算符和旋转算符的积

$$\sigma_q = \hat{I} C_{2q} = \hat{\beta} \hat{I}_{\mathrm{R}} \hat{U}_q(\pi) \hat{R}_q(\pi) \tag{4.91}$$

2. 旋转反映

如果旋转轴和对称面互相垂直，旋转之后接着进行一次反映能使物体复原，这种复合对称操作就称为旋转反映，记为 $\hat{S}_n$，有时也称为非真轴旋转。相对论的旋转反映算符写为

$$\hat{S}_n = \sigma_{\perp} C_n = \hat{\beta} \hat{I}_{\mathrm{R}} C_2 C_n = \hat{\beta} \hat{I}_{\mathrm{R}} C_m \tag{4.92}$$

式中

$$\frac{1}{m} = \frac{1}{2} + \frac{1}{n} \tag{4.93}$$

4.9 时间反演

在不考虑相对论效应时，时间反演算符 $\hat{K}_0$ 定义为：①将时间变量 t 改为 $-t$；②函数取其复共轭形式。显然，$\hat{K}_0$ 的逆算符等于它自己

$$\hat{K}_0^{-1} = \hat{K}_0, \quad \hat{K}_0 \hat{K}_0 = 1 \tag{4.94}$$

位置算符 $\hat{\boldsymbol{r}}$、动量算符 $\hat{\boldsymbol{p}}$ 和轨道角动量算符 $\hat{\boldsymbol{L}}$ 的时间反演变换分别是

$$\hat{\boldsymbol{r}}' = \hat{K}_0 \hat{\boldsymbol{r}} = \hat{\boldsymbol{r}}$$
$$\hat{\boldsymbol{p}}' = \hat{K}_0 \hat{\boldsymbol{p}} = -\hat{\boldsymbol{p}}$$
$$\hat{\boldsymbol{L}}' = \hat{K}_0 \hat{\boldsymbol{L}} = -\hat{\boldsymbol{L}} \tag{4.95}$$

1. 时间反演算符的数学性质

由于时间反演算符不只是改变时间参量的符号，还取函数的复共轭，所以它有一些特殊的性质，例如

(1) 时间反演算符是反线性算符，即

$$\hat{K}_0(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1^* \hat{K}_0\psi_1 + \lambda_2^* \hat{K}_0\psi_2 \tag{4.96}$$

(2) 时间反演算符不是厄米算符。由于 $\hat{K}_0$ 的定义中包含将函数取复共轭，所以 $\hat{K}_0$ 的转置复共轭算符 $\hat{K}_0^\dagger$ 是指将时间变量 t 改为 $-t$；显然 $\hat{K}_0^\dagger \neq \hat{K}_0$，不符合厄米算符的定义①，所以时间反演算符不是厄米算符。也就是说 $\hat{K}_0$ 虽然有本征函数，但没有物理量与之对应 (理由见附录 A)。

(3) 时间反演算符是反酉算符。虽然时间反演算符不是厄米算符，但算符 $\hat{K}_0$ 满足

$$\langle \hat{K}_0\psi | \hat{K}_0\varphi \rangle = \langle \psi | \varphi \rangle^*$$
$$|\hat{K}_0\psi| = |\psi| \tag{4.97}$$

上式符合反酉算符的定义②。因此时间反演算符有时也被称为是反酉算符。

2. 自旋 1/2 粒子体系的时间反演算符

在 4.7 节，我们将相对论空间反演算符写成自旋和位形空间的两个变换算符的乘积，然后假定 Dirac Hamilton 算符在空间反演下不变，从而求得反演算符。与此类似，考虑自旋为 1/2 的粒子，将相对论时间反演算符 $\hat{K}$ 写为 $\hat{K}_0$ 和待定算符 $\hat{a}$ 的积

$$\hat{K} = \hat{a}\hat{K}_0 \tag{4.98}$$

为确定 $\hat{a}$，将 $\hat{K}$ 作用到自旋角动量算符 $\hat{\boldsymbol{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}$，得

$$\hat{\boldsymbol{S}}' = \frac{\hbar}{2}\hat{K}\hat{\boldsymbol{\sigma}}\hat{K}^{-1} \tag{4.99}$$

由 (4.95) 式知轨道角动量在时间反演作用下改变符号，又由于自旋角动量和轨道角动量具有相同的性质，所以必有

① 厄米算符的定义见附录A

② 反酉算符的定义见附录A

$$\hat{K}\hat{\boldsymbol{\sigma}}\hat{K}^{-1} = -\hat{\boldsymbol{\sigma}} \tag{4.100}$$

将 (4.98) 式和 (2.27) 式代入，整理得

$$\begin{aligned} \hat{a}\hat{\sigma}_x\hat{a}^{-1} &= -\hat{\sigma}_x \\ -\hat{a}\hat{\sigma}_y\hat{a}^{-1} &= -\hat{\sigma}_y \\ \hat{a}\hat{\sigma}_z\hat{a}^{-1} &= -\hat{\sigma}_z \end{aligned} \tag{4.101}$$

取

$$\hat{a} = -\mathrm{i}\hat{\sigma}_y \tag{4.102}$$

则 (4.101) 式成立，于是得时间反演算符 $\hat{K}$

$$\hat{K} = -\mathrm{i}\hat{\sigma}_y\hat{K}_0 \tag{4.103}$$

且

$$\hat{K}\hat{K} = -\boldsymbol{I}_2, \quad \hat{K}^{-1} = -\hat{K} \tag{4.104}$$

由上可知，对波函数的两次时间反演将改变波函数的符号。

对多电子体系，$\hat{a}$ 应该是每个电子的 $-\mathrm{i}\hat{\sigma}_y$ 的直积，用 $\hat{U}$ 表示

$$\hat{U} = \Pi \otimes (-\mathrm{i}\hat{\sigma}_y) \tag{4.105}$$

由于 $\hat{U}$ 只作用于自旋空间，其矩阵元都是实数，所以 $\hat{U}$ 和 $\hat{K}_0$ 是对易的。如果把 (4.105) 式中的 $\hat{\sigma}_y$ 改为四分量自旋表示 $\hat{\Sigma}_y$，则得到时间反演算符的四分量表示式。

将时间反演算符作用到不含时间的 Dirac 单电子 Hamilton 算符，有

$$\hat{K}\hat{H}_{\mathrm{D}}\hat{K}^{-1} = c(-\hat{\boldsymbol{\alpha}})\cdot(-\hat{\boldsymbol{p}}) + m_0c^2\hat{K}\beta\hat{K}^{-1} + \hat{K}V\hat{K}^{-1} = c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}} + m_0c^2\beta + V \tag{4.106}$$

所以，定态单电子 Hamilton 算符在时间反演变换下不变。

将时间反演算符作用到含时间的 Dirac 方程，得

$$\begin{aligned} \hat{K}\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) &= \hat{K}\hat{H}_{\mathrm{D}}\Psi(\boldsymbol{r},t) \\ -\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{K}\Psi(\boldsymbol{r},t) &= \hat{H}_{\mathrm{D}}\hat{K}\Psi(\boldsymbol{r},t) \end{aligned} \tag{4.107}$$

令 $t' = -t$，代入上式得

$$\mathrm{i}\hbar\frac{\partial}{\partial t'}\hat{K}\Psi(\boldsymbol{r},-t') = \hat{H}_{\mathrm{D}}\hat{K}\Psi(\boldsymbol{r},-t') = \hat{H}_{\mathrm{D}}\overline{\Psi}(\boldsymbol{r},t') \tag{4.108}$$

式中，$\overline{\Psi}(\boldsymbol{r},t') = \hat{K}\Psi(\boldsymbol{r},-t')$，称为时间反演函数。由 (4.108) 式知，时间反演函数同样满足含时间的 Dirac 方程

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\overline{\Psi}(\boldsymbol{r},t) = \hat{H}_{\mathrm{D}}\overline{\Psi}(\boldsymbol{r},t) \tag{4.109}$$

也就是说，$\overline{\Psi}(\boldsymbol{r},t)$ 和 $\Psi(\boldsymbol{r},t)$ 都是 $\hat{H}_{\mathrm{D}}$ 的本征函数且本征值相同。

容易证明 $\overline{\Psi}(\boldsymbol{r},t)$ 也是归一化的，即

$$\langle\overline{\Psi}|\overline{\Psi}\rangle = \langle\hat{K}\Psi|\hat{K}\Psi\rangle = 1 \tag{4.110}$$

上式最后一个等号用到了 (4.97) 式。

我们再来计算 $\overline{\Psi}(\boldsymbol{r},t)$ 和 $\Psi(\boldsymbol{r},t)$ 的内积 $\langle\overline{\Psi}|\Psi\rangle$，将 $\hat{K}$ 作用到 $\langle\overline{\Psi}|\Psi\rangle$，得

$$\begin{aligned}\hat{K}\langle\overline{\Psi}|\Psi\rangle &= \langle\hat{K}\overline{\Psi}|\hat{K}\Psi\rangle = \langle\hat{K}_0\overline{\Psi}|(-\mathrm{i}\hat{\Sigma}_y)^\dagger(-\mathrm{i}\hat{\Sigma}_y)|\hat{K}_0\Psi\rangle = \langle\hat{K}_0\overline{\Psi}|\hat{K}_0\Psi\rangle\\ &= \langle\overline{\Psi}|\Psi\rangle^* = \langle\Psi|\hat{K}\Psi\rangle\end{aligned} \tag{4.111}$$

另外

$$\hat{K}\langle\overline{\Psi}|\Psi\rangle = \langle\hat{K}\hat{K}\Psi|\hat{K}\Psi\rangle = -\langle\Psi|\hat{K}\Psi\rangle \tag{4.112}$$

两式比较，得

$$\langle\overline{\Psi}|\Psi\rangle = 0 \tag{4.113}$$

由此我们得到，电子 (Fermi 子) 的波函数 Ψ 和其时间反演波函数 $\overline{\Psi}$ 的本征值相同且互相正交。这也证实了 Kramers 的能量简并理论[4]，即在没有外加矢量场的情况下，自旋为半整数体系的能级简并度必为偶数，且最少是双重简并。

4.10 Lorentz 变换和 Lorentz 群

在以上各节我们介绍了转动、反映、反演等空间变换的相对论表达式，本节将讨论 Lorentz 变换和空间对称变换的关系。设参考系 Σ' 以速度 u 沿另一参考系 Σ 的 x 轴运动，$\boldsymbol{r}$, $\boldsymbol{r}'$ 分别为以 Σ, Σ' 为参考系测得的四维矢量 (定义见第 1 章)

$$\boldsymbol{r} = (x,y,z,\tau), \quad \boldsymbol{r}' = (x',y',z',\tau') \tag{4.114}$$

其中 $\tau = \mathrm{i}ct, \quad \tau' = \mathrm{i}ct'$。$\boldsymbol{r}$ 和 $\boldsymbol{r}'$ 之间的 Lorentz 变换 (见 (1.13) 式) 可写为

$$\boldsymbol{r}' = \hat{\Lambda}_x(u)\boldsymbol{r} \tag{4.115}$$

式中

$$\hat{\Lambda}_x(u) = \begin{pmatrix} \gamma & 0 & 0 & \gamma\mathrm{i}u/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\mathrm{i}u/c & 0 & 0 & \gamma \end{pmatrix} \tag{4.116}$$

其中，γ 为 Lorentz 因子；$\boldsymbol{r}$ 取列向量形式。

对无限小的 Lorentz 变换，即 $u/c \ll 1$，将 γ 的级数展开式

$$\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} = 1 + \frac{u^2}{2c^2} + \cdots \tag{4.117}$$

代入 (4.116) 式并忽略 u/c 的高阶展开项，整理得

$$\hat{\varLambda}_x(u) = \boldsymbol{I}_4 - \mathrm{i}\frac{u}{c}\mathrm{Y}_x + O\left(\frac{u^2}{c^2}\right) \tag{4.118}$$

$\boldsymbol{I}_4$ 为 4×4 单位矩阵，

$$\mathrm{Y}_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{4.119}$$

(4.118) 式也可以写成矩阵指数形式

$$\hat{\varLambda}_x(u) = \mathrm{e}^{-\mathrm{i}(u/c)\mathrm{Y}_x} \tag{4.120}$$

沿 y 轴和 z 轴运动的 Lorentz 变换也容易得到相应的 $\hat{\varLambda}_y(u)$, $\hat{\varLambda}_z(u)$

$$\hat{\varLambda}_y(u) = \mathrm{e}^{-\mathrm{i}(u/c)\mathrm{Y}_y}, \quad \hat{\varLambda}_z(u) = \mathrm{e}^{-\mathrm{i}(u/c)\mathrm{Y}_z} \tag{4.121}$$

其中

$$\mathrm{Y}_y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \mathrm{Y}_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{4.122}$$

在 4.4 节，我们得到了三维空间中绕 x, y, z 轴的旋转生成元 (见 (4.29) 式和 (4.36) 式)，只需在生成元矩阵中分别增加一行和一列零元素，即得四维空间的旋转生成元

$$\mathrm{X}_x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{X}_y = \begin{pmatrix} 0 & 0 & \mathrm{i} & 0 \\ 0 & 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{X}_z = \begin{pmatrix} 0 & -\mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.123}$$

由生成元 $\mathbf{X}$ 和 $\mathbf{Y}$ 生成的群称为固有 Lorentz 群，描述了空间平动和转动的相对论变换。固有 Lorentz 群是一个连续群，群中元素可表示为 $\{\omega_1, \omega_2, \omega_3, \mathrm{i}\lambda_1, \mathrm{i}\lambda_2, \mathrm{i}\lambda_3\}$。

固有 Lorentz 群的任一无限小变换写为

$$\epsilon = \boldsymbol{I}_4 + \boldsymbol{\omega} \cdot (-\mathrm{i}\mathbf{X}) + \mathrm{i}\boldsymbol{\lambda} \cdot \mathbf{Y} \tag{4.124}$$

$\boldsymbol{\omega}$ 和 $\boldsymbol{\lambda}$ 为无限小的实矢量。$\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$，$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$。

例一 $\boldsymbol{\lambda} = (0, 0, 0)$，代入 (4.124) 式

$$\epsilon = \begin{pmatrix} 1 & -\omega_3 & \omega_2 & 0 \\ \omega_3 & 1 & -\omega_1 & 0 \\ -\omega_2 & \omega_1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.125}$$

得三维空间的转动群，ω_1, ω_2, ω_3 分别为绕 x, y, z 轴的旋转角度。

例二 $\boldsymbol{\lambda} = (u/c, 0, 0)$ 且 $u/c \ll 1$, $\boldsymbol{\omega} = (0, 0, 0)$，得沿 x 轴运动的平动 Lorentz 变换

$$\epsilon = \begin{pmatrix} 1 & 0 & 0 & \mathrm{i}u/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\mathrm{i}u/c & 0 & 0 & 1 \end{pmatrix} \tag{4.126}$$

除空间转动算符外，三维空间反演算符和时间反演算符也可以拓展到四维形式

$$\boldsymbol{I} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathrm{K}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{4.127}$$

由 $\mathbf{X}$, $\mathbf{Y}$ 和 $\mathbf{I}$ 生成的群称为正时 Lorentz 群。若再加上四维时间反演 K_0，则生成完全 Lorentz 群。Dirac 方程的对称性因而可表示为完全 Lorentz 群和 SU(2) 群的直积。

参 考 文 献

[1] Cotton F A. Chemical Applications of Group Theory. New York: John Wiley & Sons Inc, 1971.

[2] Pyykkö P, Toivonen H. Tables of representation and rotation matrices for the relativistic irreducible representations of 38 point groups. Acta Acad Aboensis Series B, 1983, 43: 1～50.

[3] Altmann S L, Herzig P. Point-Group Theory Tables. New York: Oxford University Press, 1994.

[4] Kramers H A. Théorie générale de la rotation paramagnétique dans les cristaux. Proc Royal Acad Amsterdam, 1930, 33: 959～972.

第 5 章　全电子计算法

即便是求解非相对论的 Schrödinger 方程，也只有在最简单的少数几个体系中，如势阱中的粒子、谐振子、氢原子和类氢离子、氢分子离子、刚性转动子等，才能得到精确解析解。对于稍微复杂一点的体系，如有两个及两个以上电子的原子或分子就必须用近似法来求解。与之类似，多电子体系的相对论波动方程的求解也只能用近似法来求解。就计算的对象来说，对原子和分子的所有电子作量子化学计算的方法称为全电子方法；只对价电子作严格量子化学计算的方法称为价电子 (VO) 方法，也称有效核实势法 (见第 6 章)。就求解的波函数来说，有求解四分量相对论波函数的四分量方法；也有忽略正电子，只求解电子波函数或忽略自旋，求解自旋平均波函数 (此时波函数仅含两个分量) 的二分量方法。

5.1　四分量方法

原则上，已知的非相对论计算方法如 Hartree-Fock (HF) 自洽场方法、微扰理论 (perturbation theory，PT)、组态相互作用 (configuration interaction，CI)、耦合簇 (coupled cluster，CC) 方法等均可以用于求解相对论的波动方程。另外，Dirac 方程存在正负能量解，分别对应于电子和正电子。化学工作者主要关心的是 Dirac 方程的电子解，为确保计算收敛至电子的束缚态，原有的非相对论计算过程必须加以修正。事实上，相同计算级别的相对论四分量计算程序要比非相对论程序复杂，需要更多的计算机资源。目前，精确的四分量电子相关从头算仅能计算电子结构简单的原子 (或离子) 和双原子分子。表 5.1 和表 5.2 分别列出了常见的用于原子和分子电子结构计算的相对论四分量计算程序。本节将仅介绍相对论 HF 方法和电子相关计算的一般原理，具体的计算方法，如 CI, CC 等，读者可参考表 5.1 和表 5.2 中列出的相关文献。

表 5.1　四分量的原子电子结构计算程序

方法	程序名称或主要作者	基组	参考文献
MCSCF	GRASP	数值	[1]
	MCDFGME	数值	[2]
	Ishikawa, Koc, Vilkas	GTO, STO	[3, 4]
CI	GRASP	数值	[1]
	Flambaum, Kozlov, Dzuba	数值	[5]

续表

方法	程序名称或主要作者	基组	参考文献
MRMP2	Ishikawa, Koc, Vilkas	GTO, STO	[6]
MBPT	Flambaum, Sushkov, Dzuba	数值	[7]
	Johnson, Sapirstein, Blundell	B 样条	[8]
	Lindgren, Salomonson, Mårtensson	数值	[9]
CC	Johnson, Sapirstein, Blundell	B 样条	[10]
	Salomonson, Öster	数值	[11]
FSCC	Eliav, Kaldor, Ishikawa	GTO, STO	[12]
	Das, Sahoo, Chaudhuri	GTO	[13]

注：MCSCF— 多组态自洽场；CI— 组态相互作用；MRMP2— 二级多组态 Møller-Plesset 微扰理论；MBPT— 多体微扰理论；CC— 耦合簇；FSCC—Fock 空间耦合簇；GTO—Gauss 轨道；STO—Slater 轨道。

表 5.2 四分量的分子电子结构计算程序

方法	程序名称	Hamilton 算符	参考文献
MCSCF	DIRAC	DC, DCG	[14]
CI	MOLFDIR	DC, DCG	[15]
	UTCHEM	DC	[16]
	PROPHET4R	DC	[17]
	DIRAC	DC, DCG	[14]
MP2	MOLFDIR	DC, DCG	[15]
	DIRAC	DC, DCG	[14]
MBPT	BERTHA	DC, DCB	[18]
CC	UTCHEM	DC	[16]
	DIRAC	DC, DCG	[14]
FSCC	DIRAC	DC, DCG	[14]
DFT	BDF	DC	[19]
	DIRAC	DC, DCG	[14]

注：DC—Dirac-Coulomb Hamilton 算符; DCG—DC+Gaunt; DCB—DC+Breit。

5.1.1 Hartree-Fock 方法

在非相对论的 HF 自洽场标准计算过程中，我们用变分法来求体系的近似波函数和能量是基于能量最低原理，即用任何近似波函数计算的能量一定大于或等于体系真实波函数的能量本征值。从第 2 章求解自由电子和氢原子的 Dirac 方程我们知道，相对论波动方程有正负能量解，分别对应于电子和正电子，我们寻找的电子束缚态能量是 Dirac 方程解的“激发态”，所以传统的变分法不能直接用于相

对论的 HF 计算。

1. **极小极大变分原理**

假设电子在势阱 V 中运动，将 Dirac 方程的解写成二旋量形式 $\binom{\Psi_{\rm L}}{\Psi_{\rm S}}$，则该电子的定态 Dirac 方程可写为

$$\begin{aligned} c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\rm S}+[V-E]\,\Psi_{\rm L}=0 \\ c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\rm L}+\left[V-2m_0c^2-E\right]\Psi_{\rm S}=0 \end{aligned} \tag{5.1}$$

上式采用原子单位 (定义见附录 C)。为了阅读直观，电子质量 m_0 在式中仍然保留。(5.1) 式中电子能量 E 为相对于 m_0c^2 的能量，即电子真实能量为 $E+m_0c^2$。二分量的自旋旋量 $\Psi_{\rm L}$, $\Psi_{\rm S}$ 分别表示四分量波函数的大分量和小分量。

接下来用标准的矩阵方法求解以上方程组。选取两组基函数 $\{\phi_i\}(1\leqslant i\leqslant M)$ 和 $\{\chi_j\}(1\leqslant j\leqslant N)$，将 $\Psi_{\rm L},\Psi_{\rm S}$ 分别用 $\{\phi_i\}$, $\{\chi_j\}$ 展开，于是方程 (5.1) 式的求解转化为求解 $(M+N)$ 维矩阵的本征值问题

$$\begin{pmatrix} U^{\rm LL} & cW \\ cW^{\dagger} & U^{\rm SS}-2m_0c^2\boldsymbol{I}_N \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=E\begin{pmatrix} x \\ y \end{pmatrix} \tag{5.2}$$

式中，$\boldsymbol{I}_N$ 为 $N\times N$ 单位矩阵，其他相关矩阵元为

$$U_{ij}^{\rm LL}=\langle\phi_i|V|\phi_j\rangle,\quad U_{ij}^{\rm SS}=\langle\chi_i|V|\chi_j\rangle,\quad W_{ij}=\langle\chi_i|\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}|\phi_j\rangle \tag{5.3}$$

将方程 (5.2) 式对角化得 $(M+N)$ 个本征值 E_i ，按本征值顺序排列为： $E_1\leqslant E_2\leqslant\cdots\leqslant E_{M+N}$。科学家发现，当选择合适的基函数 $\{\phi_i\}$, $\{\chi_j\}$ 时 (稍后我们再详细讨论基函数选择)，前 N 个较小本征值的能量小于 $-2m_0c^2$，对应于正电子的解[20, 21]。显然我们应该在剩下的 M 个较大本征值中寻找电子束缚态。

令 $E(M,N)$ 代表第 $(N+1)$ 个本征值，根据 Hylleraas-Undheim-MacDonald (HUM) 定理①[22, 23]，当增加一个基函数时，所得 $(M+N+1)$ 维新矩阵的本征值 $E'_{M+N+1,i}(i=1,2,\cdots,\ M+N+1)$ 和 $(M+N)$ 维原矩阵的本征值 $E_{M+N,i}(i=1,2,\cdots,\ M+N)$ 有如下关系

$$\begin{aligned} &E'_{M+N+1,1}\leqslant E_{M+N,1}\leqslant E'_{M+N+1,2}\leqslant\cdots \\ &\qquad\leqslant E'_{M+N+1,M+N}\leqslant E_{M+N,M+N}\leqslant E'_{M+N+1,M+N+1} \end{aligned} \tag{5.4}$$

由上式可见，当 N 增加而 M 不变时，$E(M,N)$ 的值变大。如果 $N\to\infty$ (即当小分量基组 $\{\chi_j\}$ 为完备基组时)，则 $E(M,\infty)$ 为 $E(M,N)$ 的上限。另外，根据 (5.4)

① HUM 定理的推导见附录E

式，$E(M+1,N) \geqslant E(M,N)$ ，也就是说，如果我们假定 $E(M,N)$ 与我们寻找的 Dirac 方程的基态能量 E_{exact} 最为接近，则 $E(M,N) \leqslant E_{\text{exact}} \leqslant E(M,\infty)$。

基于以上结论，Talman 提出了获得电子 Dirac 方程基态能量的极小极大变分原理[21] —— 首先改变小分量波函数的参数寻找第 $N+1$ 个本征值 $E(M,N)$ 的极大值，再改变大分量波函数参数极小化 $E(M,N)$，用数学公式表述就是

$$E = \min_{\Psi_{\text{L}}}\left[\max_{\Psi_{\text{S}}}\frac{\langle \Psi|\hat{H}|\Psi\rangle}{\langle \Psi|\Psi\rangle}\right] \tag{5.5}$$

2. 基函数动能平衡 (kinetic balance) 条件

由第 4 章 4.7 节知大分量波函数 Ψ_{L} 和小分量波函数 Ψ_{S} 有相反的宇称，所以两个波函数必须用不同的基函数 $\{\phi_i\}, \{\chi_j\}$ 展开，由方程 (5.1) 式中的第二个方程可得

$$c\Psi_{\text{S}} = \frac{\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}}{2m_0\left(1+\dfrac{E-V}{2m_0c^2}\right)}\Psi_{\text{L}} \tag{5.6}$$

此等式可解释为 Ψ_{S} 由 Ψ_{L} 导出。对电子束缚态，$E \ll m_0c^2$，除近核区域外，$|V(r)| \ll m_0c^2$，上式可简化为

$$\Psi_{\text{S}} \approx \frac{1}{2m_0c}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\text{L}} \tag{5.7}$$

(5.7) 式告诉我们用于波函数展开的基组 $\{\phi_i\}$ 和 $\{\chi_j\}$ 也必须满足

$$\chi_j = (\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\phi_i \tag{5.8}$$

以 Gauss 函数为例，令 $\phi = x^l y^m z^n \mathrm{e}^{-\alpha r^2}$，由于 $\hat{\boldsymbol{p}}$ 为微分算符，得

$$\begin{aligned}
\frac{\partial\phi}{\partial x} &= (1-\delta_{l0})lx^{l-1}y^m z^n\mathrm{e}^{-\alpha r^2} - 2\alpha x^{l+1}y^m z^n\mathrm{e}^{-\alpha r^2}\\
\frac{\partial\phi}{\partial y} &= (1-\delta_{m0})my^{m-1}x^l z^n\mathrm{e}^{-\alpha r^2} - 2\alpha y^{m+1}x^l z^n\mathrm{e}^{-\alpha r^2}\\
\frac{\partial\phi}{\partial z} &= (1-\delta_{n0})nz^{n-1}x^l y^m\mathrm{e}^{-\alpha r^2} - 2\alpha z^{n+1}x^l y^m\mathrm{e}^{-\alpha r^2}
\end{aligned} \tag{5.9}$$

式中，Dirac δ_{k0} 函数的定义为

$$\delta_{k0} = \begin{cases} 1 & (k=0, k=l,m,n) \\ 0 & (k\neq 0, k=l,m,n) \end{cases} \tag{5.10}$$

取 ϕ 为 p 轨道 $(l+m+n=1)$，由 (5.9) 式知必须有两个 χ (s, d) 才能满足 (5.8) 式。所以，用于小分量波函数展开的基函数的数目一般要大于大分量波函数基函数的数目。

由方程 (5.1) 式中的第一个方程，得

$$E = \frac{\langle \Psi_{\rm L}|V|\Psi_{\rm L}\rangle + \langle \Psi_{\rm L}|c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}|\Psi_{\rm S}\rangle}{\langle \Psi_{\rm L}|\Psi_{\rm L}\rangle} \tag{5.11}$$

式中，第一项为势能对能量的贡献，与 $\Psi_{\rm L}$ 有关；第二项为动能，与 $\Psi_{\rm L}$ 和 $\Psi_{\rm S}$ 均有关系，所以 (5.8) 式又称为动能平衡条件。如果此条件不能得到满足，则变分过程有可能降低 $\langle \Psi_{\rm L}|V|\Psi_{\rm L}\rangle$ 和 $\langle \Psi_{\rm L}|c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}|\Psi_{\rm S}\rangle$ 的值，导致所得能量过低，称为变分塌陷 (variational collapse)[24]。

5.1.2 电子相关能计算

1. *连续态瓦解* (continuum dissolution)

Dirac 方程负能量解的存在给电子相关能计算也带来了困难。以含有两个电子的氦原子为例，在不考虑两个电子之间相互作用的前提下，通过求解类氢原子 (核电荷数取为 2) 的 Dirac 方程得到 1s 轨道，以此轨道构成的 Slater 行列式即为氦原子的基态波函数，对应基态能量 E。另外，如果分别从正能连续态和负能连续态选取合适轨道组成 Slater 行列式也能得到相同能量本征值 E。显然，由正、负连续态混合得到的 Slater 行列式不是我们需要的解，但它们确实和电子基态波函数能量简并，当我们考虑电子–电子相互作用 (即电子相关) 时，这些行列式也会被自动计算在内，变分的结果是“连续态瓦解 (continuum dissolution)”—— 我们将得不到束缚态解。这个问题由 Brown 和 Ravenhall 首先提出[25]，因此也称为 Brown-Ravenhall 病。

2. *无虚对近似* (no-virtual pair approximation, NVPA)

为避免“连续态瓦解”，Brown 和 Ravenhall 认为必须将电子间相互作用投影到正能态，Sucher 等进一步发展了该想法 (见综述文献 [26])。

从量子电动力学的理论出发，可得到投影后的 Dirac-Coulomb-Breit Hamilton 算符[27]

$$\begin{aligned} \hat{H}^{+}_{\rm DCB} &= \sum_{i}^{n}\hat{h}_{\rm D}(i) + \sum_{i<j}^{n}\Lambda_i^{+}\hat{g}(i,j)\Lambda_j^{+} + \sum_{\lambda<\mu}^{N}\frac{Z_\lambda Z_\mu}{r_{\lambda\mu}} \\ \Lambda_i^{+} &= \sum_{m}|\psi_m(i)\rangle\langle\psi_m(i)| \end{aligned} \tag{5.12}$$

式中，$\hat{g}(i,j)$ 代表电子–电子相互作用，具体形式见 (3.51) 式。

由上式知，如果我们用同一套基函数来展开 $\hat{H}^{+}_{\rm DCB}$ 和构造投影算符 Λ_i^{+}，则 Λ_i^{+} 的作用只是将具有正能量本征值的组态视为 $\hat{H}^{+}_{\rm DCB}$ 的本征函数。在实际的相关能量计算过程中，通常的作法是直接忽略负能量解以避免“连续态瓦解”。由于 $\hat{H}^{+}_{\rm DCB}$ 避免了虚的电子–正电子对，所以又称 $\hat{H}^{+}_{\rm DCB}$ 为无对 (no-pair) DCB Hamilton 算符，所有基于 $\hat{H}^{+}_{\rm DCB}$ 的近似称为无虚对近似 (NVPA)。

关于投影算符 Λ_i^+ 中 ψ_m 的选择，常见的有以下几种：

(1) Furry 绘景[28] 取 ψ_m 为 Dirac 单电子 Hamilton 算符 (算符形式见 (3.54) 式) 的正能量本征态 —— 此方法由 Furry 提出，所以称为 Furry 绘景。常见的用于计算原子电子结构的 GRASP 程序就是采用 Furry 绘景[1]。

(2) 自由 (free) 绘景 取 ψ_m 为 Dirac 自由粒子 Hamilton 算符 (见 (2.31) 式) 的正能量本征态。研究表明，当原子序数远小于 137 时，此时的核–电子相互作用较弱，使用自由绘景能得到较好的结果[29]。

(3) 模糊 (fuzzy) 绘景[29] 选定基组后，Furry 绘景和自由绘景两种绘景下的投影算符在计算过程中是固定不变的。1981 年，Mittleman[29] 提出了动态投影算符的设想，即从分子的 HF 方程出发，由每一次的 SCF 迭代解构成投影算符，最后得到的 ψ_m 对应于 HF 收敛值。此方法已广泛用于标准的四分量计算中[4, 30]。

对多电子分子和大多数原子而言，基于无虚对近似的四分量方法是目前最精确的电子结构计算方法。20 世纪 80 年代末，Grant 等提出了多组态 Dirac-Fock (multiconfiguration Dirac-Fock，MCDF) 数值求解法[1]，用于求解多电子原子的能级和跃迁几率。MCDF 数值求解法也是第一个得到广泛应用的相对论四分量电子相关方法。至 90 年代，Vilkas 等[4] 和 Jensen 等[30] 开发了四分量的原子和分子多组态自洽场 (multiconfiguration self-consistent field，MCSCF) 方法。四分量组态相互作用 (CI) 方法也是常见的原子和分子的电子结构计算方法之一，在 GRASP [1], DIRAC [14] 等很多软件中均可使用。无论是 MCSCF 方法还是 CI 方法，均不具有体积加和性 (size-extensivity)，在用于多电子体系时需要校正。

具有体积加和性的多体微扰理论 (many-body perturbation theory，MBPT) 和耦合簇 (coupled cluster，CC) 方法也广泛用于四分量电子结构计算。与实验值相比，人们发现二级多体微扰 (MP2) 已能精确描述高价电离子中的量子电动力学 (quantum electrodynamics，QED) 效应，中性和带弱电的原子和分子则需要更高级别的相关能处理或采用多组态的多体微扰理论[31, 6]。20 世纪 90 年代出现的多参考态 Fock 空间耦合簇 (Fock-space coupled cluster，FSCC) 方法[32] 在计算重元素能级[33] 和分子结构[34] 方面与实验值符合得相当好，含重元素的分子往往由于重原子放射性而存在实验困难的问题，FSCC 方法可用于预测这些分子的电子结构。关于四分量多体微扰理论及耦合簇方法的基本原理读者可参阅 Saue 和 Visscher 撰写的综述文献 [35]。

5.1.3 密度泛函理论

1. 非相对论密度泛函基础

密度泛函理论 (density functional theory，DFT) 是原子、分子和固体的量子力学模拟中一个最重要的模拟方法。早在 20 世纪 70 年代，DFT 就已成为固体物理研

究中的标准计算方法之一。80 年代后期，在广义梯度近似 (generalized gradient approximation，GGA) 的基础上，人们开发了第二代梯度校正的交换相关泛函，DFT 从此也在分子体系的计算中流行。Hohenberg-Kohn 理论[36] 和 Kohn-Sham(KS) 动能泛函[37] 构成了 DFT 的理论基础。

假设电子在外场 $v(\boldsymbol{r})$ 中运动，根据 Hohenberg-Kohn 理论，不同的 $v(\boldsymbol{r})$ 对应不同的基态电子密度，即 $v(\boldsymbol{r})$ 是电子密度的泛函。由于电子密度也决定体系的 Hamilton 算符，所以基态 (或激发态) 能量和其他性质均是电子密度的泛函。精确的基态密度则由极小化能量泛函得到，这就是 DFT 的能量泛函变分原理。然而，能量泛函的精确求解并不比求解 Schrödinger 方程容易。为此，DFT 引入了近似能量泛函的概念，并由极小化近似能量泛函得到电子密度。这一点与基于波函数的从头算方法不同，后者是用近似方法来求解精确的 Hamilton 算符。

假定一个系统由 N 个电子组成，系统的电子密度用 $\rho(\boldsymbol{r})$ 表示，该系统的能量泛函 $E[\rho]$ 可写为

$$E[\rho]=\int v(\boldsymbol{r})\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r}+\frac{1}{2}\int\int\frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2+G[\rho] \tag{5.13}$$

式中，第一项表示电荷分布和外场的相互作用；第二项表示分布电荷的静电 Coulomb 相互作用，习惯上又称为 Hartree 能量；第三项 $G[\rho]$ 与外场 $v(\boldsymbol{r})$ 无关，当 ρ 为精确的基态电荷分布密度时，$G[\rho]$ 主要由体系的动能构成。得到 (5.13) 式的难点是无法得到 $G[\rho]$ 的精确表达式，即如何将原子或分子的动能表示为电子密度的严格泛函。迄今为止，所有努力寻找严格泛函的实验均以失败而告终。

1965 年，Kohn 和 Sham 提出将 $G[\rho]$ 写为动能隐形泛函 $T_{\mathrm{S}}[\rho]$ 和交换–相关泛函 E_{xc} 两项之和[37]。其中 $T_{\mathrm{S}}[\rho]$ 表示一个与真实体系有相同电荷密度但无电子相互作用的虚拟系统的动能，我们称此虚拟系统为真实系统的参考系。将 (5.13) 式重写为

$$E[\rho]=T_{\mathrm{S}}[\rho]+\int v(\boldsymbol{r})\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r}+\frac{1}{2}\int\int\frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2+E_{\mathrm{xc}}[\rho] \tag{5.14}$$

式中，交换–相关泛函 E_{xc} 包含两个方面的贡献，一是真实系统的电子全相互作用能和经典 Coulomb 能 (见 (5.14) 式中第三项) 的差，二是真实系统和参考系的动能差。参考系的 $T_{\mathrm{S}}[\rho]$ 可通过求解单粒子的 Schrödinger 方程得到，所得轨道称为 Kohn-Sham (KS) 轨道，这一作法和求解 Hartree-Fock 方程得到 Hartree-Fock 轨道的过程十分类似。

假定存在一组轨道 $\{\phi_i(\boldsymbol{r}), i=1,2,\cdots,N\}$，将密度 ρ 用这组轨道展开

$$\rho(\boldsymbol{r})=\sum_{i=1}^{N}\phi_i^*(\boldsymbol{r})\phi_i(\boldsymbol{r}) \tag{5.15}$$

由能量泛函最小变分原理，可得 Kohn-Sham(KS) 方程

$$\left(-\frac{1}{2}\nabla_i^2 + \upsilon^{\text{eff}}(\boldsymbol{r}_i)\right)\phi_i(\boldsymbol{r}_i) = \epsilon_i\phi_i(\boldsymbol{r}_i) \tag{5.16}$$

式中

$$\upsilon^{\text{eff}}(\boldsymbol{r}_i) = \upsilon(\boldsymbol{r}_i) + \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}_i - \boldsymbol{r}'|}\mathrm{d}\boldsymbol{r}' + \epsilon_{\text{xc}}(\boldsymbol{r}_i) \tag{5.17}$$

交换–相关能密度 ϵ_{xc} 由 E_{xc} 导出

$$\epsilon_{\text{xc}}(\boldsymbol{r}) = \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\boldsymbol{r})} \tag{5.18}$$

研究表明，当忽略 E_{xc} 时，由能量泛函最小变分原理得到的原子电子密度随 r 增大成指数衰减，正确反映了原子的壳结构。容易设想，即使简单的交换–相关近似泛函 E_{xc} 也能提高结果的精确度。最简单的近似是令 E_{xc} 是密度的局域函数，也称局域密度近似 (local density approximation，LDA)。

$$E_{\text{xc}}[\rho] = \int \rho(\boldsymbol{r})\epsilon_{\text{xc}}(\rho(\boldsymbol{r}))\mathrm{d}\boldsymbol{r} \tag{5.19}$$

式中，$\epsilon_{\text{xc}}(\rho(\boldsymbol{r}))$ 为均匀电子气 (密度为 ρ) 的交换–相关能密度。

$$\epsilon_{\text{xc}}(\rho) = \epsilon_{\text{x}}(\rho) + \epsilon_{\text{c}}(\rho) \tag{5.20}$$

式中，$\epsilon_{\text{x}}(\rho)$, $\epsilon_{\text{c}}(\rho)$ 分别为交换能密度和相关能密度，Dirac 求得均匀电子气的交换能密度为

$$\epsilon_{\text{x}}(\rho) = -\frac{3}{4}\left(\frac{3\rho}{\pi}\right)^{1/3} \tag{5.21}$$

相关能密度 $\epsilon_{\text{c}}(\rho)$ 并没有严格的函数表达形式，一般途径是由量子蒙特卡罗 (Quantum Monte Carlo，QMC) 计算得到一系列的分立值再进行函数拟合。常见的有两种不同的参数化 ϵ_{xc} 的方法[38, 39]。由于对均匀电子气的量子计算是严格的，所以 LDA 的定义也是唯一的[40]。

多数情况下可以近似认为固体中的电子是均匀分布的，符合使用 LDA 的条件，实践也证明 LDA 用于固体计算得到的结果令人十分满意，如今，LDA 已成为固体物理研究中的标准计算方法。实际上，原子或分子的电子密度分布并不均匀，不符合均匀电子气模型的要求。尽管如此，LDA 用于原子和分子体系的计算还是获得了比较好的结果[41]。

为提高 LDA 的精确度，我们可进一步要求 E_{xc} 准确描述密度变化缓慢的非均匀电子气，即梯度展开近似 (gradient expansion approximation, GEA)。研究证明梯度的低阶展开并不比 LDA 更精确[42]，这是因为真实系统的电子密度梯度往往较

大，并非密度缓慢变化的非均匀电子气。取而代之的作法是将梯度作为变量直接引入 E_{xc} 中，得到广义梯度近似 (generalized gradient approximation，GGA) 方法

$$E_{\text{xc}}[\rho] = \int F_{\text{xc}}(\rho(\boldsymbol{r}), |\nabla\rho(\boldsymbol{r})|)\mathrm{d}\boldsymbol{r} \tag{5.22}$$

考虑到转动对称性，上式假定 F_{xc} 仅和梯度的模有关。与 LDA 相比，以上函数形式并不额外增加计算量。构造泛函 F_{xc} 的方法有多种，既可从量子化学的第一原理导出，也可以采用半经验的形式[43]。GGA 方法由 Langreth 和 Mehl 两人于 1983 年首先提出，经过近 30 年的发展目前通用的 GGA 泛函不下十种，如 PW86, BP, LYP, PW91, PBE 等。遗憾的是，还没有一种泛函被公认为是最精确的。尽管如此，GGA 的出现显著提高了 DFT 的精确度，也促使 DFT 成为量子化学研究中的标准方法。

外加电磁场中的 DFT：Hohenberg-Kohn 理论并没有考虑外加磁场的影响。由 Maxwell 方程组 (见 (3.2) 式) 可知，磁场影响电流和自旋，当外磁场存在时 Hohenberg-Kohn 理论也就不再成立。针对处于强磁场的系统，Vignale 和 Rasolt 提出了流–密度和自旋–密度泛函理论[44, 45](多数情况下仅考虑磁–旋耦合)。该理论将自旋向上的电荷密度 ρ^+ 与自旋向下的电荷密度 ρ^- 视为两个独立变量，也就是说此时系统的密度由两个变量 ρ^+ 和 ρ^- 来表达。在自旋–密度泛函理论框架内，与局域密度近似 (LDA) 类似的方法称为局域自旋密度近似 (local spin density approximation，LSDA)；同样，ρ^+ 和 ρ^- 的梯度也作为两个不同变量出现在开壳层体系的 GGA 泛函中。比较体系同一电子组态的单态 ($S=0$) 和三重态 ($S=1$)，二者的电荷密度 ρ 相似，而 ρ^+, ρ^- 不同，显然此时 LSDA 对体系的描述比 LDA 更为精确。在外加磁场消失的极限条件下，LSDA 得到的能量要低于 LDA 得到的能量，这是因为 LSDA 考虑了均匀电子气的自旋极化，LDA 则忽略自旋，当密度相同时自旋极化电子气的交换能要大于非自旋极化电子气的交换能。

2. *相对论密度泛函理论*

假定电子在外场 $(\boldsymbol{a}, \phi)$ 中运动，由电磁场中单粒子的 Dirac 方程 (2.42) 式，并考虑到 (2.41) 式，得电子和外场的相互作用能 (原子单位①)

$$E_{\text{ext}} = \langle \Psi | c\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{a} | \Psi \rangle - \langle \Psi | \phi | \Psi \rangle \tag{5.23}$$

利用几率密度和几率流密度的定义 (见 (2.39) 式)，将上式重写为

$$E_{\text{ext}} = \int \hat{\boldsymbol{j}}(\boldsymbol{r}) \cdot \boldsymbol{a}(\boldsymbol{r})\mathrm{d}\boldsymbol{r} - \int \rho(\boldsymbol{r})\phi(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \tag{5.24}$$

① 原子单位定义见附录C

由上式知，电子在外场中的能量不仅和电荷密度相关，也和电流密度有关，于是，与 (5.14) 式类似，体系的相对论能量泛函 $E[\rho]$ 可写为

$$E[\rho,\boldsymbol{j}\,]=T_{\rm S}[\rho,\boldsymbol{j}\,]+\int\phi(\boldsymbol{r})\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r}+\int\hat{\boldsymbol{j}}(\boldsymbol{r})\cdot\boldsymbol{a}(\boldsymbol{r})\mathrm{d}\boldsymbol{r}$$
$$+\frac{1}{2}\iint\frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2-\frac{1}{2}\iint\frac{\boldsymbol{j}(\boldsymbol{r}_1)\boldsymbol{j}(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2+E_{\rm xc}[\rho,\boldsymbol{j}\,] \quad (5.25)$$

从能量变分原理得相对论的 Kohn-Sham 方程 (详细推导见文献 [46])

$$\left(c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}+\beta m_0c^2+v^{\rm eff}[\boldsymbol{j},\rho]\right)\mathit{\Psi}(\boldsymbol{r})=\epsilon\mathit{\Psi}(\boldsymbol{r}) \quad (5.26)$$

式中

$$v^{\rm eff}[\boldsymbol{j},\rho]=-\left[\phi(\boldsymbol{r})+\int\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}\boldsymbol{r}'+\frac{\delta E_{\rm xc}}{\delta\rho(\boldsymbol{r})}\right]$$
$$-c\hat{\boldsymbol{\alpha}}\cdot\left[\boldsymbol{a}(\boldsymbol{r})+\int\frac{\boldsymbol{j}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}\boldsymbol{r}'+\frac{\delta E_{\rm xc}}{\delta\boldsymbol{j}(\boldsymbol{r})}\right] \quad (5.27)$$

$E_{\rm xc}$ 称为相对论的交换–相关泛函, 习惯上也称上式为 Dirac-Kohn-Sham 方程。

大多数的非相对论交换–相关泛函 $E_{\rm xc}$ 是以非相对论的均匀电子气为出发点的 (见 5.1.3 节)。对均匀电子气而言，当电子密度 $\rho\approx 100\,a_0^3$ (a_0 为 Bohr 半径，此时电子的速度约为光速的 10%) 或更大时，就必须考虑相对论效应。一般来说，对密度较大的电子气，电子交换能密度 $\epsilon_{\rm x}(\rho)$ 大于电子相关能密度 $\epsilon_{\rm c}(\rho)$，是最重要的相对论能量校正来源。关于相对论的电子交换能密度的推导读者可参阅文献 [47]。相对论的相关能密度可从随机相位近似 (random phase approximation，RPA) 方法求得[48]。相对论 GGA 泛函则由半经验方法获得: 即在非相对论 GGA 交换泛函前乘以相对论校正因子，该因子来自于原子交换能的相对论校正[49]。

以上对相对论交换–相关泛函的讨论是以得到更准确的总能量为目标的，研究表明，交换–相关泛函的相对论校正对化学键能、振动频率、电子结构等化学键性质的影响很小。例如，人们发现交换能对化学键性质的影响仅为 1% 左右，与 Dirac-Kohn-Sham 方程中忽略的电磁场真空极化、电子自能等量子电动力学效应的大小在同一数量级，相关能的相对论校正的影响更小，可忽略不计。这并不是说化学键的性质不受相对论效应的影响，真实原因是大部分效应已包含在电子的相对论动能中了，即只要电子动能是由 Dirac 方程求得的，就可以忽略交换–相关泛函 $E_{\rm xc}$ 的相对论校正。目前常用的相对论 DFT 实际上是 Dirac-Kohn-Sham 方程加上非相对论 (但包括自旋) 的交换–相关泛函。需要注意的是非相对论泛函参数来自轻原子，用于重原子的计算可能会影响其精度，研究表明当 DFT 用于镧系元素计算时，会出现 4f 轨道电子布居数偏高、键长偏长的结果。

5.2 二分量方法

与非相对论的 Schrödinger 方程相比，Dirac 方程包括了小分量波函数和自旋，二者使相对论的计算量变得十分庞大。一方面，小分量波函数的存在导致了大量额外的双电子积分计算，其中主要包含了电子–电子相互作用的低阶相对论校正。鉴于化学家主要研究的是电子态 (或称正能态)，人们更愿意求解不包含小分量波函数的二分量波动方程。另一方面，在 Dirac 方程中自旋与动能算符混合，校正电子的 Coulomb 相互作用。在原子的二维旋量中，自旋与轨道角动量耦合构成了表达四分量波函数的自然基组。接下来我们讨论如何将 Dirac Hamilton 算符分解为无自旋和有自旋两项，这种作法的好处是可以在优化轨道阶段仅包含无自旋项，自旋项随后作微扰处理，最后达到降低计算量的目的。

5.2.1 自旋分离

假设单电子在势阱 V 中运动，其定态 Dirac 方程 (见 (5.1) 式) 为 (原子单位①)

$$\begin{aligned}(V-E)\,\varPsi_{\mathrm{L}}+c\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\varPsi_{\mathrm{S}}&=0\\ c\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\varPsi_{\mathrm{L}}+\left(V-E-2m_0c^2\right)\varPsi_{\mathrm{S}}&=0\end{aligned}\tag{5.28}$$

由方程组第二式可得

$$\varPsi_{\mathrm{S}}=\frac{c\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})}{E+2m_0c^2-V}\,\varPsi_{\mathrm{L}}\tag{5.29}$$

定义伪大分量波函数 $\varPhi_{\mathrm{L}}$

$$\varPhi_{\mathrm{L}}=\frac{2m_0c^2}{E+2m_0c^2-V}\,\varPsi_{\mathrm{L}}\tag{5.30}$$

代入 (5.29) 式，得

$$\varPsi_{\mathrm{S}}=\frac{\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}}{2m_0c}\,\varPhi_{\mathrm{L}}\tag{5.31}$$

将上式代入方程组 (5.28) 式，并将方程组二式左乘 $\dfrac{\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}}{2m_0c}$，得改写后的 Dirac 方程为

$$\begin{aligned}(V-E)\,\varPsi_{\mathrm{L}}+\frac{1}{2m_0}\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\varPhi_{\mathrm{L}}&=0\\ \frac{1}{2m_0}\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\varPsi_{\mathrm{L}}+\frac{1}{4m_0^2c^2}\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(V-E-2m_0c^2\right)(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\varPhi_{\mathrm{L}}&=0\end{aligned}\tag{5.32}$$

① 原子单位定义见附录C

利用 Pauli 矢量规则 (见 (2.51) 式)，有

$$\frac{1}{2m_0}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})=\frac{\hat{\boldsymbol{p}}^2}{2m_0}\boldsymbol{I}_2=\hat{T}\boldsymbol{I}_2 \tag{5.33}$$

式中，$\boldsymbol{I}_2$ 为 2×2 单位矩阵；$\hat{T}$ 为非相对论动能算符。将上式代入方程组 (5.32) 式，得

$$\begin{aligned}(V-E)\,\Psi_{\mathrm{L}}+\hat{T}\Phi_{\mathrm{L}}&=0\\ \hat{T}\Psi_{\mathrm{L}}+\left[\frac{(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,V\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})}{4m_0^2c^2}-\hat{T}-\frac{E}{2m_0c^2}\hat{T}\right]\Phi_{\mathrm{L}}&=0\end{aligned} \tag{5.34}$$

再次利用 Pauli 矢量规则 (见 (2.51) 式)，有

$$\begin{aligned}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,V\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})&=(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}V)\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\\&=\boldsymbol{I}_2\hat{\boldsymbol{p}}V\cdot\hat{\boldsymbol{p}}+\mathrm{i}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}V\times\hat{\boldsymbol{p}}\end{aligned} \tag{5.35}$$

上式中仅第二项包含自旋 $\hat{\boldsymbol{\sigma}}$，也就是说自旋仅包含在势能项中，动能项中并无自旋，联合 (5.34) 式可以看出，改写后的 Dirac 方程中自旋项仅作用于波函数 Φ_{L}。

假定势能 V 是球对称的，即 $V(\boldsymbol{r})=V(r)$，(5.35) 式中的自旋项可进一步简化为

$$\mathrm{i}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}V\times\hat{\boldsymbol{p}}=\frac{\partial V}{r\partial r}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{r}\times\hat{\boldsymbol{p}}=\frac{2}{r}\frac{\partial V}{\partial r}\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{L}} \tag{5.36}$$

上式即为我们熟悉的旋–轨耦合项。

为清楚起见，我们将 (5.34) 式重写为

$$\hat{\tilde{H}}_{\mathrm{D}}\tilde{\Psi}=E\hat{\tilde{G}}\tilde{\Psi} \tag{5.37}$$

其中 $\tilde{\Psi}$, $\hat{\tilde{G}}$, $\hat{\tilde{H}}_{\mathrm{D}}$ 分别是

$$\tilde{\Psi}=\begin{pmatrix}\Psi_{\mathrm{L}}\\ \Phi_{\mathrm{L}}\end{pmatrix},\quad \hat{\tilde{G}}=\begin{pmatrix}\boldsymbol{I}_2 & 0\\ 0 & \dfrac{\hat{T}}{2m_0c^2}\end{pmatrix} \tag{5.38}$$

$$\hat{\tilde{H}}_{\mathrm{D}}=\begin{pmatrix}V & \hat{T}\\ \hat{T} & \dfrac{\hat{\boldsymbol{p}}V\cdot\hat{\boldsymbol{p}}}{4m_0^2c^2}-\hat{T}\end{pmatrix}+\begin{pmatrix}0 & 0\\ 0 & \dfrac{\mathrm{i}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}V\times\hat{\boldsymbol{p}}}{4m_0^2c^2}\end{pmatrix} \tag{5.39}$$

(5.39) 式中第一项不包括自旋，用 $\tilde{H}^{\mathrm{sf}}$ 表示，第二项为自旋项。

5.2.2 Foldy-Wouthuysen 变换

对自由电子的 Dirac 方程，我们得到波函数的平面波解形式 (见表 2.1)，两个正能解的大分量和小分量的模之比约为 a^2，而 a^2 是量级为 v^2/c^2 的一个小量 (见

(2.75) 式)。此外，氢原子的相对论波函数也有类似的特点，即正能解的大分量远大于小分量 (见图 2.1)。产生大小分量混合的原因是 Dirac Hamilton 算符中存在奇、偶两种类型的矩阵，所谓奇矩阵就是 4×4 矩阵中有两个 2×2 的块对角元为零的非零矩阵，如 $\hat{\boldsymbol{\alpha}}$ 就是奇矩阵 (也称奇算符)；偶矩阵则相反，是两个 2×2 的块对角元，而非对角元为零的矩阵，$\hat{\beta}$ 就是偶矩阵 (也称偶算符)。容易证明，奇算符和奇算符相乘得到的新算符必为偶算符，偶算符与偶算符相乘仍为偶算符，而奇算符与偶算符相乘得到的新算符仍为奇算符。

以电磁场中单电子的 Dirac Hamilton 算符(见 (2.42) 式) 为例，Hamilton 算符的主要项 $\hat{\beta}m_0c^2$ 以及 $q\phi$ 是偶矩阵，但另一项 $c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{\pi}}$ 却是奇矩阵，后者与前者之比的量级约为 v/c 。

设想对 Hamilton 算符作一个酉变换，将电子的 Hamilton 算符变成偶矩阵，那么所有 Hamilton 算符的正能解将不包括小分量，这样就可以忽略负能解了。这个变换最早由 Foldy 和 Wouthuysen 提出，称为 Foldy-Wouthuysen 变换，简称 FW 变换。

1. **自由电子 Dirac Hamilton 算符的 FW 变换**

为便于讨论，我们取自由电子的 Hamilton 算符在动量空间中的表示 (原子单位①)

$$\hat{H} = c\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p} + \hat{\beta}m_0c^2 \tag{5.40}$$

式中，$\boldsymbol{p}$ 即为动量空间中的动量算符，注意此时动量算符不再是微分算符 (见附录 G)。

现在寻找一个酉变换 $\mathrm{e}^{\mathrm{i}\hat{S}}$，将 (5.40) 式变成偶矩阵。为简化讨论，仅考虑定态 Dirac 方程的情形，此时 $\hat{S}$ 与时间无关，由 (4.7) 式得变换后的 Hamilton 算符 $\hat{H}_1$ 与原 Hamilton 算符有下列关系

$$\hat{H}_1 = \mathrm{e}^{\mathrm{i}\hat{S}}\hat{H}\mathrm{e}^{-\mathrm{i}\hat{S}} \tag{5.41}$$

取 $\hat{S}$ 为如下形式

$$\hat{S} = -\left(\frac{\mathrm{i}}{2m_0c}\right)\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}f\left(\frac{p}{m_0c}\right) \tag{5.42}$$

式中，函数 $f\left(\dfrac{p}{m_0c}\right)$ 待定。考虑到 $\hat{S}$ 与算符 $\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}$ 对易，将 (5.41) 式重写为

$$\begin{aligned}\hat{H}_1 &= \mathrm{e}^{\mathrm{i}\hat{S}}\left(c\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p} + \hat{\beta}m_0c^2\right)\mathrm{e}^{-\mathrm{i}\hat{S}}\\ &= \mathrm{e}^{\mathrm{i}\hat{S}}\hat{\beta}\left(c\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p} + m_0c^2\right)\mathrm{e}^{-\mathrm{i}\hat{S}}\\ &= \mathrm{e}^{\mathrm{i}\hat{S}}\hat{\beta}\mathrm{e}^{-\mathrm{i}\hat{S}}\hat{\beta}\left(c\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p} + \hat{\beta}m_0c^2\right)\end{aligned} \tag{5.43}$$

① 原子单位定义见附录C

利用指数函数 $\mathrm{e}^{-\mathrm{i}\hat{S}}$ 的级数展开，并考虑到 (5.42) 式及

$$\hat{\beta}\left(\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}\right)^j=(-1)^j\left(\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}\right)^j\hat{\beta} \tag{5.44}$$

得

$$\begin{aligned}\hat{\beta}\mathrm{e}^{-\mathrm{i}\hat{S}}&=\hat{\beta}\sum_{j=0}^{\infty}\left(\frac{-1}{2m_0c}\right)^j\frac{\left(\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}\right)^j}{j!}f^j\\&=\sum_{j=0}^{\infty}\left(\frac{1}{2m_0c}\right)^j\frac{\left(\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}\right)^j}{j!}f^j\hat{\beta}=\mathrm{e}^{\mathrm{i}\hat{S}}\hat{\beta}\end{aligned} \tag{5.45}$$

代入 (5.43) 式，有

$$\hat{H}_1=\mathrm{e}^{2\mathrm{i}\hat{S}}\left(c\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}+\hat{\beta}m_0c^2\right) \tag{5.46}$$

由于 $\mathrm{e}^{2\mathrm{i}\hat{S}}=\cos(2\hat{S})+\mathrm{i}\sin(2\hat{S})$，并考虑到三角函数 $\cos(2\hat{S})$, $\sin(2\hat{S})$ 的级数展开式

$$\begin{aligned}\cos(2\hat{S})&=1-\frac{(p/m_0c)^2f^2}{2!}+\frac{(p/m_0c)^4f^4}{4!}-\cdots=\cos\left(\frac{p}{m_0c}f\right)\\\sin(2\hat{S})&=\frac{\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}}{p}\left(\frac{p}{m_0c}f-\frac{(p/m_0c)^3f^3}{3!}+\cdots\right)=\frac{\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}}{p}\sin\left(\frac{p}{m_0c}f\right)\end{aligned} \tag{5.47}$$

上式利用了 (5.42) 式和 $(\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p})(\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p})=p^2$。
将 (5.46) 式重写为

$$\begin{aligned}\hat{H}_1&=\left[\cos\left(\frac{p}{m_0c}f\right)+\frac{\hat{\beta}\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}}{p}\sin\left(\frac{p}{m_0c}f\right)\right]\left[c\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}+\hat{\beta}m_0c^2\right]\\&=\hat{\beta}\cos\left(\frac{p}{m_0c}f\right)\left[m_0c^2+cp\tan\left(\frac{p}{m_0c}f\right)\right]\\&\quad+c\hat{\boldsymbol{\alpha}}\cdot\boldsymbol{p}\cos\left(\frac{p}{m_0c}f\right)\left[1-\frac{m_0c}{p}\tan\left(\frac{p}{m_0c}f\right)\right]\end{aligned} \tag{5.48}$$

为消除 $\hat{H}_1$ 中的奇矩阵，可令上式第二个等式中的第二项为零，以求得待定函数 f

$$f=\frac{m_0c}{p}\arctan\left(\frac{p}{m_0c}\right) \tag{5.49}$$

将 (5.49) 式代入 (5.48) 式，考虑到三角函数关系式 $\arctan x=\arccos\left(\dfrac{1}{\sqrt{1+x^2}}\right)$，得

$$\begin{aligned}\hat{H}_1&=\hat{\beta}\frac{m_0c}{\sqrt{m_0^2c^2+p^2}}\left(m_0c^2+\frac{p^2}{m_0}\right)\\&=\hat{\beta}E_p\end{aligned} \tag{5.50}$$

式中，$E_p=\sqrt{p^2c^2+m_0^2c^4}$。

由 (5.50) 式可见，新 Hamilton 算符 $\hat{H}_1$ 中只有偶算符，可以将自由电子 Dirac 方程的正负能态完全分离。

实际上，酉变换 $\mathrm{e}^{\mathrm{i}\hat{S}}$ 的选择并不唯一，新定义酉变换 $\hat{U}$

$$\hat{U}=\frac{\hat{\beta}\left(c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}\right)+m_0c^2+E_p}{\sqrt{2E_p\left(m_0c^2+E_p\right)}} \tag{5.51}$$

读者可自己证明，$\hat{U}$ 也能将自由电子的 Dirac Hamilton 算符变换成 (5.50) 式。

2. 氢原子 Dirac Hamilton 算符的 FW 变换

无外加电磁场时，氢原子的 Dirac Hamilton 算符可写为 (原子单位①)

$$\hat{H}=\hat{\beta}m_0c^2+\hat{\varepsilon}+\hat{O} \tag{5.52}$$

式中，$\hat{\varepsilon}$, $\hat{O}$ 分别为 Hamilton 算符中的偶算符和奇算符

$$\hat{\varepsilon}=V,\quad \hat{O}=c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}} \tag{5.53}$$

Hamilton 算符中各项的量级可作如下估计，假定 V 为氢原子基态 Bohr 轨道上的电势，则

$$V=\frac{e}{a_0\sqrt{4\pi\epsilon_0}},\quad u=\frac{\hbar}{m_0a_0} \tag{5.54}$$

式中，u 表示 Bohr 轨道上电子的速度；$a_0,\epsilon_0,\hbar$ 分别表示 Bohr 半径、真空介电常量和约化 Planck 常量。采用原子单位，可得 Hamilton 算符中 $\hat{\varepsilon}$, $\hat{O}$ 的量级

$$\begin{aligned}\hat{\varepsilon}&=V\sim c^0\\ \hat{O}&=c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}\approx cm_0u\sim c^1\end{aligned} \tag{5.55}$$

接下来我们要寻找一个酉变换 $\mathrm{e}^{\mathrm{i}\hat{S}}$，将 (5.52) 式变成偶算符。与 (5.42) 式类似，取

$$\hat{S}=-\frac{\mathrm{i}}{2m_0c^2}\hat{\beta}\hat{O} \tag{5.56}$$

由于 $\hat{O}$ 为 c^1 量级，所以 $\hat{S}$ 为 c^{-1} 量级。利用附录 A 中的 (A.18) 式，将 (5.41) 式重写为

$$\begin{aligned}\hat{H}_1&=\mathrm{e}^{\mathrm{i}\hat{S}}\hat{H}\mathrm{e}^{-\mathrm{i}\hat{S}}\\ &=\sum_{j=0}^{\infty}\frac{(\mathrm{i})^j}{j!}\left[\hat{S}^{(j)},\hat{H}\right]\\ &=\hat{H}+\mathrm{i}\left[\hat{S},\hat{H}\right]+\frac{\mathrm{i}^2}{2!}\left[\hat{S},\left[\hat{S},\hat{H}\right]\right]+\cdots\end{aligned} \tag{5.57}$$

① 原子单位定义见附录C

由于 $\hat{S}$ 是一个量级为 c^{-1} 的小量，上式展开项的量级随 j 增大而逐渐降低。

下面给出 $j=1,2,3,4$ 时各项的计算结果，此时展开式相当于精确到 c^{-2}

$$
\begin{aligned}
\mathrm{i}\left[\hat{S},\hat{H}\right] &= -\hat{O}+\frac{\hat{\beta}}{2m_0c^2}\left[\hat{O},\hat{\varepsilon}\right]+\frac{\hat{\beta}}{m_0c^2}\hat{O}^2\\
\frac{\mathrm{i}^2}{2!}\left[\hat{S},\left[\hat{S},\hat{H}\right]\right] &= -\frac{\hat{\beta}}{2m_0c^2}\hat{O}^2-\frac{1}{2m_0^2c^4}\hat{O}^3-\frac{1}{8m_0^2c^4}\left[\hat{O},\left[\hat{O},\hat{\varepsilon}\right]\right]\\
\frac{\mathrm{i}^3}{3!}\left[\hat{S},\left[\hat{S},\left[\hat{S},\hat{H}\right]\right]\right] &= \frac{1}{6m_0^2c^4}\hat{O}^3-\frac{\hat{\beta}}{6m_0^3c^6}\hat{O}^4-\frac{\hat{\beta}}{48m_0^3c^6}\left[\hat{O},\left[\hat{O},\left[\hat{O},\hat{\varepsilon}\right]\right]\right]\\
\frac{\mathrm{i}^4}{4!}\left[\hat{S},\left[\hat{S},\left[\hat{S},\left[\hat{S},\hat{H}\right]\right]\right]\right] &= \frac{\hat{\beta}}{24m_0^3c^6}\hat{O}^4+\cdots
\end{aligned}
\tag{5.58}
$$

将以上结果代回到 (5.57) 式，整理得 $\hat{H}_1$ 为

$$
\hat{H}_1=\hat{\beta}m_0c^2+\hat{\varepsilon}_1+\hat{O}_1 \tag{5.59}
$$

偶算符 $\hat{\varepsilon}_1$ 和奇算符 $\hat{O}_1$ 分别为

$$
\begin{aligned}
\hat{\varepsilon}_1 &= \hat{\varepsilon}+\frac{\hat{\beta}}{2m_0c^2}\hat{O}^2-\frac{\hat{\beta}}{8m_0^3c^6}\hat{O}^4-\frac{1}{8m_0^2c^4}\left[\hat{O},\left[\hat{O},\hat{\epsilon}\right]\right]\\
\hat{O}_1 &= \frac{\hat{\beta}}{2m_0c^2}\left[\hat{O},\hat{\epsilon}\right]-\frac{1}{3m_0^2c^4}\hat{O}^3-\frac{\hat{\beta}}{48m_0^3c^6}\left[\hat{O},\left[\hat{O},\left[\hat{O},\hat{\epsilon}\right]\right]\right]
\end{aligned}
\tag{5.60}
$$

此时奇算符 $\hat{O}_1$ 的量级为 c^{-1}，小于未变换前的奇算符量级 c。为使 $\hat{H}_1$ 中 $\hat{O}_1$ 的量级进一步减少，需对 $\hat{H}_1$ 再作一次酉变换。

取

$$
\hat{S}'=-\frac{\mathrm{i}}{2m_0c^2}\hat{\beta}\hat{O}_1 \tag{5.61}
$$

得二次变换后的 Hamilton 算符为

$$
\begin{aligned}
\hat{H}_2 &= \mathrm{e}^{\mathrm{i}\hat{S}'}\hat{H}'\mathrm{e}^{-\mathrm{i}\hat{S}'}\\
&= \hat{\beta}m_0c^2+\hat{\varepsilon}_1+\frac{\beta}{2m_0c^2}\left[\hat{O}_1,\varepsilon_1\right]-\frac{1}{3m_0^2c^4}\hat{O}_1^3+\cdots
\end{aligned}
\tag{5.62}
$$

式中，奇算符的量级已进一步降低到 c^{-3}，可以略去，偶算符中量级在 c^{-4} 的小项也一并忽略。

将 $\hat{\varepsilon}_1$(见 (5.60) 式) 代入上式，最后我们得到仅含偶算符的氢原子的 Hamilton 算符为

$$
\hat{H}=\hat{\beta}m_0c^2+\hat{\varepsilon}+\frac{\hat{\beta}}{2m_0c^2}\hat{O}^2-\frac{\hat{\beta}}{8m_0^3c^6}\hat{O}^4-\frac{1}{8m_0^2c^4}\left[\hat{O},\left[\hat{O},\hat{\varepsilon}\right]\right] \tag{5.63}
$$

该算符的最小量级为 c^{-2}。

下面计算 (5.63) 式。将 (5.53) 式代入并利用 Pauli 矢量规则 (见 (2.51) 式)，有

$$
\begin{aligned}
\frac{\hat{\beta}}{2m_0c^2}\hat{O}^2 &= \frac{\hat{\beta}\hat{\boldsymbol{p}}^2}{2m_0} \\
\frac{\hat{\beta}}{8m_0^3c^6}\hat{O}^4 &= \frac{\hat{\beta}\hat{\boldsymbol{p}}^4}{8m_0^3c^2}
\end{aligned}
\tag{5.64}
$$

现在计算 $\left[\hat{O},\hat{\epsilon}\right]$，将 $\left[\hat{O},\hat{\varepsilon}\right]$ 作用到任一函数 ψ

$$
\begin{aligned}
\left[\hat{O},\hat{\varepsilon}\right]\psi &= c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}V\psi - cV\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}\psi \\
&= \psi\left(-\mathrm{i}c\hat{\boldsymbol{\alpha}}\cdot\nabla V\right)
\end{aligned}
\tag{5.65}
$$

比较等式两边，得

$$
\left[\hat{O},\hat{\varepsilon}\right] = -\mathrm{i}c\hat{\boldsymbol{\alpha}}\cdot\nabla V \tag{5.66}
$$

其次计算

$$
\begin{aligned}
&-\frac{1}{8m_0^2c^4}\left[\hat{O},\left[\hat{O},\hat{\varepsilon}\right]\right] \\
&= \frac{\mathrm{i}}{8m_0^2c^2}\left[(\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}})(\hat{\boldsymbol{\alpha}}\cdot\nabla V) - (\hat{\boldsymbol{\alpha}}\cdot\nabla V)(\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}})\right] \\
&= \frac{\mathrm{i}}{8m_0^2c^2}\left[(\hat{\boldsymbol{p}}\cdot\nabla V - \nabla V\cdot\hat{\boldsymbol{p}}) + \mathrm{i}\hat{\boldsymbol{\Sigma}}\cdot(\hat{\boldsymbol{p}}\times\nabla V - \nabla V\times\hat{\boldsymbol{p}})\right]
\end{aligned}
\tag{5.67}
$$

式中，$\hat{\boldsymbol{\Sigma}}$ 为 Pauli 矩阵的四维形式 (见 (2.59) 式)。

上式方括号内的第一项即为 $[\hat{\boldsymbol{p}},\nabla V]$，采用与求 $\left[\hat{O},\hat{\epsilon}\right]$ 相似的方法 (见 (5.65) 式)，可得

$$
[\hat{\boldsymbol{p}},\nabla V] = -\mathrm{i}\nabla^2 V \tag{5.68}
$$

再来求方括号内的另一项 $\hat{\boldsymbol{p}}\times\nabla V - \nabla V\times\hat{\boldsymbol{p}}$，将其作用到任一函数 ψ，得

$$
\begin{aligned}
(\hat{\boldsymbol{p}}\times\nabla V - \nabla V\times\hat{\boldsymbol{p}})\psi &= -\mathrm{i}\nabla\times(\nabla V\psi) + \mathrm{i}\nabla V\times\nabla\psi \\
&= -\mathrm{i}\psi\nabla\times\nabla V + \mathrm{i}\nabla V\times\nabla\psi + \mathrm{i}\nabla V\times\nabla\psi \\
&= -\mathrm{i}\psi\frac{\partial\boldsymbol{B}}{\partial t} - 2\nabla V\times\hat{\boldsymbol{p}}\psi \\
&= -2\nabla V\times\hat{\boldsymbol{p}}\psi
\end{aligned}
\tag{5.69}
$$

上式推导从第一步到第二步用到了矢量关系 (见附录 B 的 (B.8) 式)

$$
\nabla\times(\boldsymbol{a}\psi) = \psi(\nabla\times\boldsymbol{a}) - \boldsymbol{a}\times(\nabla\psi) \tag{5.70}
$$

式中，$\boldsymbol{a}$ 为任意矢量。(5.69) 式推导从第二步到第三步用到了电场定义 $\boldsymbol{E}=-\nabla V$ 及电场 $\boldsymbol{E}$ 、磁场 $\boldsymbol{B}$ 的关系式，即 Maxwell 方程组的第二式(见 (3.2) 式)；最后一步则使用了定态条件 $(\partial \boldsymbol{B}/\partial t=0)$。

比较 (5.69) 式两边，得

$$\hat{\boldsymbol{p}}\times\nabla V-\nabla V\times\hat{\boldsymbol{p}}=-2\nabla V\times\hat{\boldsymbol{p}} \tag{5.71}$$

将上式和 (5.68) 式代入 (5.67) 式，有

$$\begin{aligned}-\frac{1}{8m_0^2c^4}\left[\hat{O},\left[\hat{O},\hat{\varepsilon}\right]\right]&=\frac{1}{8m_0^2c^2}\nabla^2V+\frac{1}{4m_0^2c^2}\hat{\boldsymbol{\Sigma}}\cdot(\nabla V\times\hat{\boldsymbol{p}})\\&=\frac{1}{8m_0^2c^2}\nabla^2V+\frac{1}{2m_0^2c^2}\cdot\frac{1}{r}\frac{\partial V}{\partial r}\left(\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{L}}\right)\end{aligned} \tag{5.72}$$

式中，$\hat{\boldsymbol{S}}$ 为自旋的四维表示

$$\hat{\boldsymbol{S}}=\frac{1}{2}\hat{\boldsymbol{\Sigma}} \tag{5.73}$$

最后得正负能态分离在 c^{-4} 量级内的 Hamilton 算符为

$$\hat{H}=\hat{\beta}\left(m_0c^2+\frac{\hat{\boldsymbol{p}}^2}{2m_0}-\frac{\hat{\boldsymbol{p}}^4}{8m_0^3c^2}\right)+V+\frac{1}{8m_0^2c^2}\nabla^2V+\frac{1}{2m_0^2c^2}\cdot\frac{1}{r}\frac{\partial V}{\partial r}\left(\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{L}}\right) \tag{5.74}$$

将上式中的 β 取为 1，(5.73) 式中的 $\hat{\boldsymbol{\Sigma}}$ 取为 $\hat{\boldsymbol{\sigma}}$，即可得出正能态空间中的 Hamilton 算符

$$\hat{H}_{\text{Pauli}}=m_0c^2+\frac{\hat{\boldsymbol{p}}^2}{2m_0}-\frac{\hat{\boldsymbol{p}}^4}{8m_0^3c^2}+V+\frac{1}{8m_0^2c^2}\nabla^2V+\frac{1}{2m_0^2c^2}\cdot\frac{1}{r}\frac{\partial V}{\partial r}\left(\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{L}}\right) \tag{5.75}$$

上式即为 Pauli Hamilton 算符。

与非相对论 Hamilton 算符相比，(5.75) 式中第一项为电子的静止能量，是一个常数，可以改变能量的起点而消去。第二、四项为电子的非相对论动能和势能，可合写为 $\hat{H}_{\text{NR}}$；第三项为电子动能项的高一级校正，又称质量–速度 (mass-velocity) 项，用 $\hat{H}_{\text{MV}}$ 表示；第五项称为 Darwin 项，这一项是对势能项 V 的修正，用 $\hat{H}_{\text{D}}^1$ 表示；第六项则是我们熟知的单电子旋–轨耦合，用 $\hat{H}_{\text{SO}}^1$ 表示。

采用点电荷核模型，此时

$$\begin{aligned}V&=-Z/r\\\frac{\partial V}{\partial r}&=\frac{Z}{r^3}\boldsymbol{r}\\\nabla^2V&=\nabla\cdot\boldsymbol{E}=4Z\pi\delta(\boldsymbol{r})\end{aligned} \tag{5.76}$$

式中，Z 为核电荷数；$\delta(\boldsymbol{r})$ 为 Dirac δ 函数。式中第三式的第二个等号用到了 Maxwell 方程组 (3.2) 式。将上式代入 (5.75) 式并略去静止能量，得类氢原子 Pauli Hamilton 算符

$$\hat{H}_{\text{Pauli}} = \hat{H}_{\text{NR}} + \hat{H}_{\text{MV}} + \hat{H}_{\text{D}}^1 + \hat{H}_{\text{SO}}^1 \tag{5.77}$$

其中

$$\begin{aligned}
\hat{H}_{\text{NR}} &= \frac{\hat{\boldsymbol{p}}^2}{2m_0} - Z/r \\
\hat{H}_{\text{MV}} &= -\frac{\hat{\boldsymbol{p}}^4}{8m_0^3c^2} \\
\hat{H}_{\text{D}}^1 &= \frac{Z\pi}{2m_0^2c^2}\delta(\boldsymbol{r}) \\
\hat{H}_{\text{SO}}^1 &= \frac{Z}{2m_0^2c^2r^3}\hat{\boldsymbol{S}} \cdot \hat{\boldsymbol{L}}
\end{aligned} \tag{5.78}$$

3. 分子的 Breit-Pauli(BP) Hamilton 算符

分子是多电子体系，为得到二分量 Hamilton 算符，我们还需要对电子–电子及原子–原子相互作用项进行 FW 变换。在忽略核自旋及核运动的条件下，在 FW 变换下可近似认为原子–原子 Coulomb 相互作用形式保持不变。将 FW 变换作用到电子–电子 Coulomb 相互作用 $\hat{g}_{\text{c}}$，并略去量级小于 c^{-2} 的展开项，有

$$\begin{aligned}
\hat{g}'_{\text{c}} &= \mathrm{e}^{\mathrm{i}\hat{S}}\hat{g}_{\text{c}}\mathrm{e}^{-\mathrm{i}\hat{S}} \\
&\approx \sum_{i<j}\left\{\frac{1}{r_{ij}} + \frac{1}{8m_0^2c^2}\left[\left(\nabla_i^2\frac{1}{r_{ij}}\right) + \left(\nabla_j^2\frac{1}{r_{ij}}\right)\right]\right. \\
&\quad \left. - \frac{1}{4m_0^2c^2r_{ij}^3}\left[\hat{\boldsymbol{\sigma}}_i \cdot \boldsymbol{r}_{ij} \times \hat{\boldsymbol{p}}_i + \hat{\boldsymbol{\sigma}}_j \cdot \boldsymbol{r}_{ij} \times \hat{\boldsymbol{p}}_j\right]\right\} \\
&= \sum_{i<j}\left[\frac{1}{r_{ij}} - \frac{\pi}{m_0^2c^2}\delta(\boldsymbol{r}_{ij}) - \frac{1}{m_0^2c^2r_{ij}^3}\hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{L}}_{ij}\right]
\end{aligned} \tag{5.79}$$

式中，求和指标 i, j 表示对所有电子求和。上式最后一个等式包含三项，第一项仍为 Coulomb 相互作用，第二项为双电子 Darwin 项，用 $\hat{H}_{\text{D}}^2$ 表示，最后一项为双电子旋–轨耦合算符，用 $\hat{H}_{\text{SO}}^2$ 表示，表示第 i 个电子的自旋与该电子围绕第 j 个电子的轨道角动量的相互作用，由于此时自旋与动量属于同一个电子，所以又称 $\hat{H}_{\text{SO}}^2$ 为自旋–相同的–轨道 (spin-same-orbit) 相互作用。

联合 (5.77) 式，分子的 Pauli Hamilton 算符可写为

$$\hat{H}_{\text{Pauli}} = \hat{H}_{\text{NR}} + \hat{H}_{\text{MV}} + \hat{H}_{\text{D}}^1 + \hat{H}_{\text{D}}^2 + \hat{H}_{\text{SO}}^1 + \hat{H}_{\text{SO}}^2 \tag{5.80}$$

式中

$$
\begin{aligned}
\hat{H}_{\mathrm{NR}} &= \sum_i \frac{\hat{\boldsymbol{p}}_i^2}{2m_0} - \sum_{i\lambda} \frac{Z_\lambda}{r_{i\lambda}} + \sum_{i<j} \frac{1}{r_{ij}} + \sum_{\mu<\lambda} Z_\lambda Z_\mu / R_{\lambda\mu} \\
\hat{H}_{\mathrm{MV}} &= -\frac{1}{8m_0^3c^2} \sum_i \hat{\boldsymbol{p}}_i^4 \\
\hat{H}_{\mathrm{D}}^1 &= \sum_i \frac{Z\pi}{2m_0^2c^2} \delta(\boldsymbol{r}_i) \\
\hat{H}_{\mathrm{D}}^2 &= -\sum_{i<j} \frac{\pi}{m_0^2c^2} \delta(\boldsymbol{r}_{ij}) \\
\hat{H}_{\mathrm{SO}}^1 &= \sum_{i\lambda} \frac{Z_\lambda}{2m_0^2c^2r_{i\lambda}^3} \hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{L}}_{i\lambda} \\
\hat{H}_{\mathrm{SO}}^2 &= \sum_{i<j} -\frac{1}{m_0^2c^2r_{ij}^3} \hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{L}}_{ij}
\end{aligned}
\tag{5.81}
$$

式中，求和指标 i,j 表示对所有电子求和；λ,μ 表示对所有原子核求和。

接下来考虑电子–电子间 Coulomb-Breit 相互作用(见 (3.51) 式)，对其实施 FW 变换，得

$$
\begin{aligned}
\hat{g}'_{\mathrm{CB}} = \hat{g}'_{\mathrm{C}} + \sum_{i<j} \Bigg[&-\frac{1}{2m_0^2c^2} \left(\frac{\hat{\boldsymbol{p}}_i \cdot \hat{\boldsymbol{p}}_j}{r_{ij}} + \frac{\boldsymbol{r}_{ij} \left(\boldsymbol{r}_{ij} \cdot \hat{\boldsymbol{p}}_j \right) \cdot \hat{\boldsymbol{p}}_i}{r_{ij}^3} \right) \\
&- \frac{2}{m_0^2c^2r_{ij}^3} \hat{\boldsymbol{S}}_j \cdot \boldsymbol{r}_{ij} \times \hat{\boldsymbol{p}}_i - \frac{8\pi}{3m_0^2c^2} \left(\hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j \right) \delta(\boldsymbol{r}_{ij}) \\
&+ \frac{1}{m_0^2c^2} \left(\frac{\hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j}{r_{ij}^3} - \frac{3 \left(\hat{\boldsymbol{S}}_i \cdot \boldsymbol{r}_{ij} \right) \left(\hat{\boldsymbol{S}}_j \cdot \boldsymbol{r}_{ij} \right)}{r_{ij}^5} \right) \Bigg]
\end{aligned}
\tag{5.82}
$$

式中，方括号内第一项描述轨道–轨道相互作用，用 $\hat{H}_{\mathrm{OO}}$ 表示；第二项描述第 j 个电子的自旋和第 i 个电子相对于第 j 个电子的轨道角动量之间的相互作用，此时自旋和轨道角动量属于不同的电子，所以又称为自旋–其他的–轨道 (spin-other-orbit) 相互作用，用 $\hat{H}_{\mathrm{SO}}^3$ 表示；方括号内余下两项为电子的自旋–自旋相互作用，用 $\hat{H}_{\mathrm{SS}}$ 表示。最后我们得到分子的 BP Hamilton 算符为

$$
\begin{aligned}
\hat{H}_{\mathrm{BP}} &= \hat{H}_{\mathrm{Pauli}} + \hat{H}_{\mathrm{Breit}} \\
&= \hat{H}_{\mathrm{NR}} + \hat{H}_{\mathrm{MV}} + \hat{H}_{\mathrm{D}}^1 + \hat{H}_{\mathrm{D}}^2 + \hat{H}_{\mathrm{SO}}^1 + \hat{H}_{\mathrm{SO}}^2 + \hat{H}_{\mathrm{SO}}^3 + \hat{H}_{\mathrm{OO}} + \hat{H}_{\mathrm{SS}}
\end{aligned}
\tag{5.83}
$$

式中，$\hat{H}_{\mathrm{SO}}^1$, $\hat{H}_{\mathrm{SO}}^2$, $\hat{H}_{\mathrm{SO}}^3$ 三项之和合称为 Breit-Pauli 旋–轨耦合算符。

4. Pauli Hamilton 算符存在的问题

在实际计算中，人们发现

(1) Pauli Hamilton 算符存在奇点。观察 Pauli Hamilton 算符的单电子旋–轨耦合算符

$$\hat{H}^1_{\mathrm{SO}} = \frac{Z}{2m_0^2c^2r^3}\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{L}} \tag{5.84}$$

当 $r\to 0$ 时上式趋于无穷大。双电子旋–轨耦合算符 $\hat{H}^2_{\mathrm{SO}}$，$\hat{H}^3_{\mathrm{SO}}$ 也存在类似问题。使用有限核电荷密度分布模型 (见 2.4 节) 可避免奇点问题。

(2) Pauli Hamilton 算符中动量的四次方算符 $\hat{\boldsymbol{p}}^4$ 在 Hilbert 空间中并不是一个好算符，它可以产生非平方可积的函数[50]，所以 Pauli Hamilton 算符通常用于微扰理论，此时的零级 Hamilton 算符为非相对论 Hamilton 算符 $\hat{H}_{\mathrm{NR}}$，(5.83) 式中所有其他项为微扰项。

(3) Pauli Hamilton 算符用于重元素计算时产生较大误差。在寻找 FW 变换时，我们估计氢原子 Hamilton 算符中奇算符的量级为 c^1(见 (5.55) 式)，电子速度近似为 1 a.u.①，远小于光速 c (137 a.u.)。在此近似下二次 FW 变换后的 Pauli Hamilton 算符精确到 c^{-4} 量级。可是电子的速度并不总是远小于光速，用 Bohr 模型估计原子中电子的速度

$$v\sim\frac{Z_{\mathrm{eff}}}{n} \tag{5.85}$$

式中，n 为轨道量子数；Z_{eff} 为有效核电荷。显然，当 Z_{eff} 较大 n 较小时，如重元素的 1s 轨道，其速度 ($v_{1\mathrm{s}}\approx Z$) 已不再满足远小于光速的条件。此时二次 FW 变换已不能满足要求，需要多次近似。对轻的分子体系，研究表明 Pauli Hamilton 算符已经能相当精确地描述相对论效应[51]。

(4) Pauli Hamilton 算符对基组的收缩系数十分敏感，这是因为相对论效应主要来自内层轨道 (近核区域的电子速度快)。基组收缩对非相对论波函数影响很小，而内层轨道系数的微小改变可导致大的相对论校正。

5.2.3 Douglas-Kroll-Hess Hamilton 算符

FW 变换得到的二分量 Pauli Hamilton 算符包含动量的四次方算符 $\hat{\boldsymbol{p}}^4$，不具备变分稳定性，所以 Pauli Hamilton 算符多用于微扰计算。接下来我们介绍由 Douglas 和 Kroll 提出的另一种酉变换 —— Douglas-Kroll 变换，简称为 DK 变换。变换后得到的 Hamilton 算符是以势能为展开参数的，所得二分量 Hamilton 算符可以用于变分计算[52]。

1. DK 变换

将自由电子的 FW 变换 (见 (5.51) 式) 重写为

$$\hat{U}_0 = \hat{A}\boldsymbol{I}_4 + \hat{\beta}\hat{A}\hat{R} \tag{5.86}$$

① 1 a.u. $=2.188\times10^6$ m· s^{-1}

式中，$\boldsymbol{I}_4$ 为 4×4 单位矩阵。$\hat{A}$, $\hat{R}$ 分别为

$$\hat{A}=\sqrt{\frac{\hat{E}_p+mc^2}{2\hat{E}_p}},\quad \hat{R}=\frac{c\hat{\boldsymbol{\alpha}}\cdot\hat{\boldsymbol{p}}}{\hat{E}_p+m_0c^2} \tag{5.87}$$

根据 (5.55) 式容易估计出 $\hat{A}$, $\hat{R}$ 的量级分别为 c^0 和 c^{-1}。将 $\hat{U}_0$ 作用到氢原子的 Dirac Hamilton 算符 (5.52) 式，得

$$\begin{aligned}\hat{H}_1&=\hat{U}_0\hat{H}\hat{U}_0^{-1}\\&=\hat{\beta}E_p+\hat{\varepsilon}_1+\hat{O}_1\end{aligned} \tag{5.88}$$

式中，$\hat{\varepsilon}_1$, $\hat{O}_1$ 为含有势能 V 的偶算符和奇算符

$$\begin{aligned}\hat{\varepsilon}_1&=\hat{A}(V+\hat{R}V\hat{R})\hat{A}\\\hat{O}_1&=-\beta\hat{A}[V,\hat{R}]\hat{A}\end{aligned} \tag{5.89}$$

此时 $\hat{O}_1$ 的量级是 c^{-1}。

为进一步降低奇算符的量级，Douglas 和 Kroll 定义新的酉变换

$$\hat{U}_1=\sqrt{1+\hat{W}_1^2}+\hat{W}_1 \tag{5.90}$$

式中，$\hat{W}_1$ 待定。

容易证明，为使 $\hat{U}_1$ 为酉算符，$\hat{W}_1$ 必为反厄米算符①，即当 $\hat{W}_1^\dagger=-\hat{W}_1$ 时

$$\begin{aligned}\hat{U}_1^\dagger\hat{U}_1&=\left(\sqrt{1+\hat{W}_1^2}-\hat{W}_1\right)\left(\sqrt{1+\hat{W}_1^2}+\hat{W}_1\right)\\&=1\end{aligned} \tag{5.91}$$

将 $\hat{U}_1$ 作用到 $\hat{H}_1$，并利用 $\sqrt{1+\hat{W}_1^2}=1+\hat{W}_1+(1/2)\hat{W}_1^2+\cdots$，得

$$\begin{aligned}\hat{H}_2&=\hat{U}_1\hat{H}_1\hat{U}_1^{-1}\\&=\left(\sqrt{1+\hat{W}_1^2}+\hat{W}_1\right)\hat{H}_1\left(\sqrt{1+\hat{W}_1^2}-\hat{W}_1\right)\\&=\hat{H}_1+\left[\hat{W}_1,\hat{H}_1\right]+\frac{1}{2}\left[\hat{W}_1,\left[\hat{W}_1,\hat{H}_1\right]\right]+\cdots\\&=\hat{\beta}E_p+\hat{\varepsilon}_1+\hat{O}_1-\left[\hat{\beta}E_p,\hat{W}_1\right]+\left[\hat{W}_1,\hat{\varepsilon}_1\right]+\left[\hat{W}_1,\hat{O}_1\right]\\&\quad+\frac{1}{2}\left[\hat{W}_1,\left[\hat{W}_1,\hat{\beta}E_p\right]\right]+\cdots\end{aligned} \tag{5.92}$$

① 反厄米算符的定义见附录A

为消除奇算符 $\hat{O}_1$，令

$$\left[\hat{\beta}E_p, \hat{W}_1\right] = \hat{O}_1 \tag{5.93}$$

上式左边 $\hat{\beta}E_p$ 为偶算符，右边为奇算符，为使上式成立，得 $\hat{W}_1$ 必为奇算符。

将 (5.93) 式左边展开，整理得

$$\begin{aligned}\left[\hat{\beta}E_p, \hat{W}_1\right] &= \hat{\beta}E_p\hat{W}_1 - \hat{W}_1\hat{\beta}E_p \\ &= E_p\hat{W}_1 + \hat{W}_1E_p = \hat{\beta}\hat{O}_1\end{aligned} \tag{5.94}$$

上式利用了 $\hat{\beta}^2 = 1$ 及奇算符 $\hat{W}_1$ 必与 $\hat{\beta}$ 反对易的性质。Douglas 和 Kroll 发现，如果使用算符在动量空间①的表示，则很容易从上式求得 $\hat{W}_1$。

在动量空间中动量算符 $\hat{\boldsymbol{p}} = \boldsymbol{p}$，不再是微分算符 $-\mathrm{i}\hbar\nabla$。动量空间中的一般算符 $\hat{g}$ 表示为积分算符形式 (证明见附录G)

$$\hat{g}f(\boldsymbol{p}) = \int \hat{g}(\boldsymbol{p}, \boldsymbol{p}')f(\boldsymbol{p}')\mathrm{d}\boldsymbol{p}' \tag{5.95}$$

式中，$f(\boldsymbol{p})$ 为动量空间的任一函数；$\hat{g}(\boldsymbol{p}, \boldsymbol{p}')$ 称为算符 $\hat{g}$ 的内核。Douglas 和 Kroll 求得动量空间中算符 $\hat{W}_1$ 的内核为

$$\hat{W}_1(\boldsymbol{p}, \boldsymbol{p}') = \frac{\hat{\beta}}{E_p + E'_p}\hat{O}_1(\boldsymbol{p}, \boldsymbol{p}') \tag{5.96}$$

由 $\hat{O}_1$ 的量级 (c^{-1}) 可估计出 $\hat{W}_1$ 的量级为 c^{-3}。下面我们证明上式确为 (5.94) 式的解。

将 (5.94) 式左边的算符作用于任一函数 $f(\boldsymbol{p})$，并将 (5.96) 式代入，得

$$\begin{aligned}\left[E_p\hat{W}_1 + \hat{W}_1E_p\right]f(\boldsymbol{p}) &= E_p\int \frac{\hat{\beta}}{E_p + E'_p}\hat{O}_1(\boldsymbol{p}, \boldsymbol{p}')f(\boldsymbol{p}')\mathrm{d}\boldsymbol{p}' \\ &\quad + \int \frac{\hat{\beta}}{E_p + E'_p}\hat{O}_1(\boldsymbol{p}, \boldsymbol{p}')E'_pf(\boldsymbol{p}')\mathrm{d}\boldsymbol{p}' \\ &= \int \beta\hat{O}_1(\boldsymbol{p}, \boldsymbol{p}')f(\boldsymbol{p}')\mathrm{d}\boldsymbol{p}'\end{aligned} \tag{5.97}$$

显然，上式即为将 (5.94) 式右边的算符作用于任一函数 $f(\boldsymbol{p})$ 的结果，证明完毕。

将 $\hat{O}_1$ 的表达式 (见 (5.89) 式) 代入 (5.96) 式，得

$$\hat{W}_1(\boldsymbol{p}, \boldsymbol{p}') = \frac{\hat{A}\hat{R}V(\boldsymbol{p}, \boldsymbol{p}')\hat{A}' - \hat{A}V(\boldsymbol{p}, \boldsymbol{p}')\hat{R}'\hat{A}'}{E_p + E'_p} \tag{5.98}$$

① 关于动量空间的介绍见附录G

在动量空间中，势能 V 也表示成积分算符。以原子 (核电荷数为 Z) 为例，V 的内核可写为 (原子单位，推导过程见附录 G)

$$V(\boldsymbol{p},\boldsymbol{p}') = -Z\left(\frac{2}{\pi}\right)^{1/2}\cdot\frac{1}{|\boldsymbol{p}-\boldsymbol{p}'|^2} \tag{5.99}$$

按 (5.93) 式消除奇算符后，二次酉变换后的 Hamilton 算符 (见 (5.92) 式) 重写为

$$\begin{aligned}\hat{H}_2 &= \hat{U}_1\hat{U}_0\hat{H}\hat{U}_0^{-1}\hat{U}_1^{-1}\\ &= \hat{\beta}E_p + \hat{\varepsilon}_1 + \left[\hat{W}_1,\hat{\varepsilon}_1\right] + \left[\hat{W}_1,\hat{O}_1\right] + \frac{1}{2}\left[\hat{W}_1,\left[\hat{W}_1,\hat{\beta}E_p\right]\right] + \cdots\\ &= \hat{\beta}E_p + \hat{\varepsilon}_1 + \left[\hat{W}_1,\hat{\varepsilon}_1\right] - \frac{1}{2}\left[\hat{W}_1,\left[\hat{W}_1,\hat{\beta}E_p\right]\right]\end{aligned} \tag{5.100}$$

上式中第三项为奇算符，用 $\hat{O}_2$ 表示，量级为 c^{-5}，可以略去。第四项为偶算符，量级为 c^{-4}，也可以略去。

最后我们得到仅含偶算符的氢原子 Hamilton 算符为

$$\hat{H}_{\mathrm{DK2}} = \hat{\beta}E_p + \hat{\varepsilon}_1 \tag{5.101}$$

该算符准确到 c^{-4} 量级。

DK 变换可一直进行，直至无穷。一般地，n 次变换后的 Hamilton 算符可写为

$$\begin{aligned}\hat{H}_{\mathrm{DK}n} &= \hat{U}_n\hat{U}_{n-1}\cdots\hat{U}_0\hat{H}\hat{U}_0^{-1}\cdots\hat{U}_{n-1}^{-1}\hat{U}_n^{-1} = \sum_{k=0}^{n}\hat{\varepsilon}_k\\ \hat{U}_n &= \left(1+\hat{W}_n^2\right)^{1/2} + \hat{W}_n\end{aligned} \tag{5.102}$$

积分算符 $\hat{W}_n$ 的核定义为

$$\hat{W}_n(\boldsymbol{p},\boldsymbol{p}') = \frac{\hat{\beta}}{E_p+E_p'}\hat{O}_n(\boldsymbol{p},\boldsymbol{p}') \tag{5.103}$$

式中，奇算符 $\hat{O}_n$ 量级是 V^n，由第 $(n-1)$ 次 DK 变换得到。关于 DK 高级变换的详细讨论，读者请参阅 Nakajima 和 Hirao 的综述文献 [56]。

比较 FW 变换和 DK 变换我们发现，前者以动量为展开参数，展开式中包含动量算符的高次幂，后者以 V 为展开参数，避免了动量算符的高次幂出现，所以 DK 变换得到的 Hamilton 算符具有变分稳定性。

2. Douglas-Kroll-Hess 近似

下面将一级 DK 变换 $\hat{U}_0$(见 (5.86) 式) 作用到电子–电子 Coulomb 相互作用项 $\hat{g}_{ij}$，得

$$\begin{aligned}\hat{g}'_{ij} &= \hat{U}_{0,i}\hat{U}_{0,j}\hat{g}_{ij}\hat{U}_{0,j}^{-1}\hat{U}_{0,i}^{-1}\\ &= \hat{A}_i\hat{A}_j\left(\hat{g}_{ij} + \hat{R}_i\hat{g}_{ij}\hat{R}_i + \hat{R}_j\hat{g}_{ij}\hat{R}_j + \hat{R}_i\hat{R}_j\hat{g}_{ij}\hat{R}_j\hat{R}_i\right)\hat{A}_j\hat{A}_i\\ &= \hat{g}_{ij} + \left[\hat{U}_{0,i},\hat{g}_{ij}\right]\hat{U}_{0,i}^{-1} + \left[\hat{U}_{0,j},\hat{g}_{ij}\right]\hat{U}_{0,j}^{-1} + \left[\hat{U}_{0,i},\left[\hat{U}_{0,j},\hat{g}_{ij}\right]\right]\hat{U}_{0,j}^{-1}\hat{U}_{0,i}^{-1}\end{aligned} \tag{5.104}$$

式中，$U_{0,i}, \hat{U}_{0,j}$ 分别表示作用于第 i, j 个电子的一级 DK 变换。原则上，我们还需对上式进行二次酉变换，以获得变换后的电子–电子相互作用。然而，由于电子–电子相互作用远小于电子与核的相互作用 (相差约为 Z 倍，Z 为核电荷数)，人们发现一级 DK 变换对计算结果的影响非常小，其量级约为单电子算符的 5 级 DK 变换。因此，在不考虑自旋时，通常只取 (5.104) 式中的第一项，即 Coulomb 算符 $\hat{g}_{ij} = \dfrac{1}{r_{ij}}$，其余项均忽略不计。该近似由 Hess 首先提出，称为 Douglas-Kroll-Hess (DKH) 近似。联合 (5.101) 式，我们得到多电子体系的 DKH Hamilton 算符

$$\hat{H}_{\rm DKH} = \sum_i \hat{\beta} E_p(i) + \sum_i \hat{\varepsilon}_1(i) + \sum_{i<j} \frac{1}{r_{ij}} \tag{5.105}$$

式中，求和符号 i 表示对所有电子求和。其中

$$\begin{aligned}
\hat{\varepsilon}_1(i) &= \hat{A}_p(i)\left(V(i) + \hat{R}_p(i)V(i)\hat{R}_p(i)\right)\hat{A}_p(i) \\
E_p(i) &= \sqrt{\boldsymbol{p}^2(i)c^2 + m_0^2c^4} \\
\hat{A}_p(i) &= \sqrt{\frac{E_p(i) + m_0c^2}{2E_p(i)}} \\
\hat{R}_p(i) &= \frac{c\boldsymbol{p}(i)}{E_p(i) + m_0c^2}
\end{aligned} \tag{5.106}$$

DKH Hamilton 算符准确到 c^{-4} 量级。上式不含自旋，实际上是一个一分量的 Hamilton 算符。

3. 旋–轨耦合

将 DK 变换 $\hat{U}_0$(见 (5.86) 式) 作用到电子–电子相对论相互作用项 ——Coulomb-Breit 相互作用 $\hat{g}_{\rm CB}(i,j)$(见 (3.51) 式)，利用 Pauli 矢量规则分离自旋，得[53]

$$\begin{aligned}
\hat{H}_{\rm SO} = &\sum_{i\lambda} Z_\lambda f_1(\boldsymbol{p}_i)\hat{\boldsymbol{\sigma}}_i \cdot \left(\frac{\boldsymbol{r}_{i\lambda}}{r_{i\lambda}^3} \times \boldsymbol{p}_i\right) f_1(\boldsymbol{p}_i) \\
&-2\sum_{i<j} f_2(\boldsymbol{p}_i, \boldsymbol{p}_j)\hat{\boldsymbol{\sigma}}_i \cdot \left(\frac{\boldsymbol{r}_{ij}}{r_{ij}^3} \times \boldsymbol{p}_i\right) f_2(\boldsymbol{p}_i, \boldsymbol{p}_j) \\
&-4\sum_{i<j} f_2(\boldsymbol{p}_i, \boldsymbol{p}_j)\hat{\boldsymbol{\sigma}}_j \cdot \left(\frac{\boldsymbol{r}_{ij}}{r_{ij}^3} \times \boldsymbol{p}_i\right) f_2(\boldsymbol{p}_j, \boldsymbol{p}_i)
\end{aligned} \tag{5.107}$$

式中，λ, i 分别表示对所有原子和电子求和；Z_λ 表示第 λ 个原子的原子序数。式中第一项是单电子旋–轨耦合算符，用 $\hat{H}_{\rm SO}^1$ 表示，余下的两项代表双电子旋–轨耦合算符，用 $\hat{H}_{\rm SO}^2$ 表示。

$$f_1(\boldsymbol{p}_i) = \frac{cA_i}{E_i + m_0c^2}, \quad f_2(\boldsymbol{p}_i, \boldsymbol{p}_j) = \frac{cA_iA_j}{E_i + m_0c^2} \tag{5.108}$$

考虑能量的级数展开 $E=\sqrt{p^2c^2+m_0^2c^4}=m_0c^2+\dfrac{p^2}{2m_0}+\cdots$，如果我们只保留展开式第一项，即 $E\approx m_0c^2$，由系数 A 的定义 (见 (5.87) 式)，有 $A=\sqrt{(E+m_0c^2)/2E}\approx 1$，代入 (5.108) 式，得

$$f_1(\boldsymbol{p}_i)=f_2(\boldsymbol{p}_i,\boldsymbol{p}_j)\approx\frac{1}{2m_0c} \tag{5.109}$$

算符 $\hat{H}_{\mathrm{SO}}^1$, $\hat{H}_{\mathrm{SO}}^2$ 可简化为

$$\begin{aligned}\hat{H}_{\mathrm{SO}}^1&=\sum_{i\lambda}Z_\lambda\frac{1}{2m_0c^2}\hat{\boldsymbol{S}}_i\cdot\left(\frac{\boldsymbol{r}_{i\lambda}}{r_{i\lambda}^3}\times\boldsymbol{p}_i\right)\\ \hat{H}_{\mathrm{SO}}^2&=\sum_{i<j}\frac{-1}{m_0c^2}\left(\hat{\boldsymbol{S}}_i+2\hat{\boldsymbol{S}}_j\right)\cdot\left(\frac{\boldsymbol{r}_{ij}}{r_{ij}^3}\times\boldsymbol{p}_i\right)\end{aligned} \tag{5.110}$$

上式即为 Breit-Pauli (BP) 旋–轨耦合算符 (见 (5.83) 式中的 $H_{\mathrm{SO}}^1+\hat{H}_{\mathrm{SO}}^2+\hat{H}_{\mathrm{SO}}^3$)。

4. *旋–轨耦合的平均场近似*

观察 (5.110) 式我们发现，双电子旋–轨耦合算符 $\hat{H}_{\mathrm{SO}}^2$ 的符号与单电子旋–轨耦合算符 $\hat{H}_{\mathrm{SO}}^1$ 相反，研究表明 $\hat{H}_{\mathrm{SO}}^2$ 对自旋–轨道分裂的贡献可达 20%~50%，在实际计算中不能忽略。另外，双电子的自旋–轨道积分的数目比双电子 Coulomb 积分的数目要高出一个数量级，为简化计算，Hess 等[54] 首先提出了平均场旋–轨耦合的方法，采用该方法后，我们只需要计算单电子的自旋–轨道积分。

下面我们以原子的自旋–轨道组态相互作用 (SOCI) 方法为例来说明平均场近似的应用。将体系的波函数 $\varPsi$ 写为一组 Slater 行列式的线性组合，对其中的一个行列式 $\varPhi$，组成该行列式 $\varPhi$ 的自旋轨道用 m 表示，则关于 $\varPhi$ 的单电子旋–轨耦合算符的矩阵元可写为

$$\langle\varPhi|\hat{H}_{\mathrm{SO}}^1|\varPhi\rangle=\sum_m n(\varPhi,m)\langle m|\hat{h}_{\mathrm{SO}}^1|m\rangle \tag{5.111}$$

式中

$$\hat{H}_{\mathrm{SO}}^1=\sum_i\hat{h}_{\mathrm{SO}}^1(i)=\sum_i Z\frac{1}{2m_0c^2}\hat{\boldsymbol{S}}_i\cdot\left(\frac{\boldsymbol{r}_i}{r_{i^3}}\times\boldsymbol{p}_i\right) \tag{5.112}$$

(5.111) 式中，当 m 为占据轨道时，$n(\varPhi,m)$ 等于 1；当 m 为空轨道时，$n(\varPhi,m)$ 等于 0。由 (5.112) 式知单电子旋–轨耦合算符在空间反演下是反对称的，所以对整个空间的积分 (即 (5.111) 式) 必为零。

考虑单激发态 $\varPhi_m^p$，此时 $\varPhi$ 中的一个轨道 m 由另一轨道 p 代替，相关矩阵元为

$$\langle\varPhi|\hat{H}_{\mathrm{SO}}^1|\varPhi_m^p\rangle=\langle m|\hat{h}_{\mathrm{SO}}^1|p\rangle \tag{5.113}$$

接着来看双电子算符，为能使用积分计算的 Slater-Condon 规则，先将双电子算符写成关于求和指标 i,j 互换对称的形式

$$\hat{H}_{\mathrm{SO}}^2=\sum_{i<j}\hat{h}_{\mathrm{SO}}^2(i,j)=\frac{1}{2}\sum_{i<j}\left[\hat{h}_{\mathrm{SO}}^2(i,j)+\hat{h}_{\mathrm{SO}}^2(j,i)\right] \tag{5.114}$$

与单电子旋–轨耦合算符一样，双电子旋–轨耦合算符也具有空间反对称性，所以对角矩阵元为零

$$\langle\varPhi|\hat{H}_{\mathrm{SO}}^2|\varPhi\rangle=0 \tag{5.115}$$

由 Slater-Condon 规则，得单激发态 $\varPhi_m^p$ 与 $\varPhi$ 之间的矩阵元为

$$\begin{aligned}&\langle\varPhi|\frac{1}{2}\sum_{i<j}\left[\hat{h}_{\mathrm{SO}}^2(i,j)+\hat{h}_{\mathrm{SO}}^2(j,i)\right]|\varPhi_m^p\rangle\\=&\frac{1}{2}\sum_{l}n(\varPhi,l)\Big[\langle ml|\hat{h}_{\mathrm{SO}}^2(1,2)|pl\rangle+\langle ml|\hat{h}_{\mathrm{SO}}^2(2,1)|pl\rangle-\langle ml|\hat{h}_{\mathrm{SO}}^2(1,2)|lp\rangle\\&-\langle ml|\hat{h}_{\mathrm{SO}}^2(2,1)|lp\rangle\Big]\end{aligned} \tag{5.116}$$

式中，求和符号 l 表示对所有不包括 m 和 p 的占据轨道求和。注意到方括号内的第二个积分为零 (利用旋–轨耦合算符的空间反对称性)，上式可进一步简化为

$$\begin{aligned}&\langle\varPhi|\frac{1}{2}\sum_{i<j}\left[\hat{h}_{\mathrm{SO}}^2(i,j)+\hat{h}_{\mathrm{SO}}^2(j,i)\right]|\varPhi_m^p\rangle\\=&\frac{1}{2}\sum_{l}n(\varPhi,l)\Big[\langle ml|\hat{h}_{\mathrm{SO}}^2(1,2)|pl\rangle-\langle ml|\hat{h}_{\mathrm{SO}}^2(1,2)|lp\rangle-\langle lm|\hat{h}_{\mathrm{SO}}^2(1,2)|pl\rangle\Big]\end{aligned} \tag{5.117}$$

联合 (5.113) 式，得单激发 $m\to p$ 条件下的旋–轨耦合矩阵元为

$$\begin{aligned}\langle\varPhi|\hat{H}_{\mathrm{SO}}|\varPhi_m^p\rangle=&\langle m|\hat{h}_{\mathrm{SO}}^1|p\rangle+\frac{1}{2}\sum_{l}n(\varPhi,l)\Big[\langle ml|\hat{h}_{\mathrm{SO}}^2(1,2)|pl\rangle\\&-\langle ml|\hat{h}_{\mathrm{SO}}^2(1,2)|lp\rangle-\langle lm|\hat{h}_{\mathrm{SO}}^2(1,2)|pl\rangle\Big]\end{aligned} \tag{5.118}$$

上式可看成是价电子 (电荷分布由轨道 m,p 决定) 在由其他轨道 l 的电子产生的场中运动。假定原子外有一个未成对的 d 电子，d 的自旋轨道共有 10 个，且这 10 个轨道均不是 (5.118) 式中的 m 或 p。为简化积分计算，我们可以用平均化的轨道 $\tilde{v}$ 代替所有的 d 轨道，它们的电子占据数均为 1/10 —— 此即原子的平均场 (mean field) 近似。于是，(5.118) 式可近似为

$$\begin{aligned}\langle \Phi|\hat{H}_{\text{SO}}^{\text{mean field}}|\Phi_m^p\rangle = &\langle m|\hat{h}_{\text{SO}}^1|p\rangle + \frac{1}{2}\sum_i\Big[\langle mi|\hat{h}_{\text{SO}}^2(1,2)|pi\rangle - \langle mi|\hat{h}_{\text{SO}}^2(1,2)|ip\rangle \\ &-\langle im|\hat{h}_{\text{SO}}^2(1,2)|pi\rangle\Big] + \frac{1}{2}\sum_v n(v)\Big[\langle m\tilde{v}|\hat{h}_{\text{SO}}^2(1,2)|p\tilde{v}\rangle \\ &-\langle m\tilde{v}|\hat{h}_{\text{SO}}^2(1,2)|\tilde{v}p\rangle - \langle \tilde{v}m|\hat{h}_{\text{SO}}^2(1,2)|p\tilde{v}\rangle\Big]\end{aligned} \tag{5.119}$$

式中，i 对所有的闭壳层求和；v 对所有的开壳层求和；$n(v)$ 代表开壳层 v 的电子平均占据数；$\tilde{v}$ 表示平均化后的开壳层轨道。

读者可能已经注意到，与 (5.118) 式相比，我们需计算的自旋–轨道积分的数目并没有减少，因此 (5.119) 式还需作进一步近似。由 (5.110) 式知，旋–轨耦合算符和 r^3 成反比，于是我们有理由认为自旋–轨道积分的主要贡献来自近原子核区域 (r 较小)，因此在计算中可以忽略多中心的双电子自旋–轨道积分 —— 很多研究表明此近似带来的误差并不大。

最后我们提一下行列式 Φ 中两个轨道被激发的情形，如双激发态 $\Phi_{m,n}^{p,q}$，此时 Φ 中的两个轨道 m,n 分别由 p,q 代替，易得单电子旋–轨耦合算符的矩阵元为零。对双电子旋–轨耦合算符 $\hat{H}_{\text{SO}}^2$，由 Slater-Condon 规则，得双激发态 $\Phi_{m,n}^{p,q}$ 与 Φ 之间的矩阵元为

$$\begin{aligned}&\langle \Phi|\frac{1}{2}\sum_{i<j}\Big[\hat{h}_{\text{SO}}^2(i,j) + \hat{h}_{\text{SO}}^2(j,i)\Big]|\Phi_{m,n}^{p,q}\rangle \\ =&\frac{1}{2}\Big[\langle mn|\hat{h}_{\text{SO}}^2(1,2)|pq\rangle + \langle nm|\hat{h}_{\text{SO}}^2(1,2)|qp\rangle - \langle mn|\hat{h}_{\text{SO}}^2(1,2)|qp\rangle \\ &-\langle nm|\hat{h}_{\text{SO}}^2(1,2)|pq\rangle\Big]\end{aligned} \tag{5.120}$$

考虑到双电子旋–轨耦合算符的数量级约为 $1/c^2$(见 (5.110) 式)，实际计算中往往不考虑双激发的情形。实践经验也表明，以上积分对最后结果的贡献并不大。

5. DKH 求解

为求解 DKH Hamilton 算符，我们需要在动量空间计算所有积分。DKH Hamilton 算符 (见 (5.105) 式和 (5.106) 式) 中的算符分两类，一类为非相对论 Hamilton 算符中已存在的算符，即核–电子相互作用算符 V 及电子–电子相互作用算符 $1/r_{ij}$，涉及此类算符的积分可直接用已有的非相对论量子化学程序计算，在此不用考虑；另一类为非相对论 Hamilton 算符中没有的算符，如 $E_p, \hat{A}_p, (\boldsymbol{p}\cdot V\boldsymbol{p})$ 等。其中 $E_p, \hat{A}_p$ 均为 $\boldsymbol{p}^2$ 的函数，当选定基组 $\{\phi_i\}(i=1,2,\cdots,n)$ 时， Hess 等提出了如下步骤以求得动量空间中任一以 $\boldsymbol{p}^2$ 为变量的函数 $f(\boldsymbol{p}^2)$ 的矩阵表示

$$M_{\mu\nu}[f(\boldsymbol{p}^2)] \equiv \langle \phi_\mu|f(\boldsymbol{p}^2)|\phi_\nu\rangle \quad (\mu,\nu=1,2,\cdots,n) \tag{5.121}$$

(1) 将算符 $\boldsymbol{p}^2$ 在给定基组下的矩阵表示 $\mathbf{M}[\boldsymbol{p}^2]$ 变换到正交归一化基组 $\{X_i\}(i=1,2,\cdots,n)$，得新矩阵表达式 $\tilde{\mathbf{M}}[\boldsymbol{p}^2]$ 为

$$\tilde{\mathrm{M}}[\boldsymbol{p}^2]=\mathrm{X}^\dagger\mathrm{M}[\boldsymbol{p}^2]\mathrm{X} \tag{5.122}$$

式中, X 满足

$$1=\mathrm{X}^\dagger\mathrm{SX},\quad S_{\mu\nu}=\langle\phi_\mu|\phi_\nu\rangle \tag{5.123}$$

(2) 将矩阵 $\tilde{\mathrm{M}}[\boldsymbol{p}^2]$ 对角化，得

$$\tilde{\mathrm{M}}[\boldsymbol{p}^2]\mathrm{U}=\mathrm{U}\lambda \tag{5.124}$$

式中，U 为酉矩阵；λ 为对角矩阵，对角元用 $\lambda_i(i=1,2,\cdots,n)$ 表示。将上式右乘 $\mathrm{U}^\dagger$，有

$$\tilde{\mathrm{M}}[\boldsymbol{p}^2]=\mathrm{U}\lambda\mathrm{U}^\dagger \tag{5.125}$$

(3) 函数 $f(\boldsymbol{p}^2)$ 在正交归一化基组 $\{X_i\}(i=1,2,\cdots,n)$ 下的表达矩阵为

$$\tilde{\mathrm{M}}[f(\boldsymbol{p}^2)]=\mathrm{Uf}(\lambda)\mathrm{U}^\dagger \tag{5.126}$$

式中，$\mathrm{f}(\lambda)$ 为对角矩阵，对角元为 $f(\lambda_i)(i=1,2,\cdots,n)$。

(4) 将 $\tilde{\mathrm{M}}[f(\boldsymbol{p}^2)]$ 变换回原始基组 $\{\phi_i\}(i=1,2,\cdots,n)$，得

$$\begin{aligned}\mathrm{M}[f(\boldsymbol{p}^2)]&=(\mathrm{X}^\dagger)^{-1}\tilde{\mathrm{M}}[f(\boldsymbol{p}^2)]\mathrm{X}^{-1}\\&=\mathrm{SXUf}(\lambda)(\mathrm{SXU})^\dagger\end{aligned} \tag{5.127}$$

$\mathrm{M}[f(\boldsymbol{p}^2)]$ 即为给定基组 $\{\phi_i\}(i=1,2,\cdots,n)$ 下 $f(\boldsymbol{p}^2)$ 的矩阵表示。
以上方法的好处是可以直接利用非相对论量子化学程序中关于非相对论动能 $T=\boldsymbol{p}^2/2m_0$ 的计算程序。

DKH Hamilton 算符 (见 (5.105) 式) 中另一类算符 $\boldsymbol{p}\cdot V\boldsymbol{p}$ 的矩阵元计算可利用如下关系[55]

$$\langle\phi_\mu|\boldsymbol{p}\cdot V\boldsymbol{p}|\phi_\nu\rangle=\langle\nabla\phi_\mu|V|\nabla\phi_\nu\rangle\quad(\mu,\nu=1,2,\cdots,n) \tag{5.128}$$

这样，原有的非相对论计算程序只需略作改动便可用于 DKH 计算，DKH 方法于是很快成为量子化学程序中的标准方法，并得到了广泛的应用。虽然上述展开只有在所选基组为完备基组时才严格成立，研究表明非收缩的基函数就能得到令人满意的结果[56]。

对周期表中第一周期以后的元素研究表明，如果在 DKH 计算中使用非相对论的收缩基组将可能导致较大的误差，所以在求解 DKH Hamilton 算符时应使用配

套的相对论基组。表 5.3 列出了常见的 DKH 基组，对应于二 (或三) 次 DK 变换得到的 Hamilton 算符 (分别用 DK2, DK3 表示)。表 5.3 中，文献 [57]~[61] 中的基组是选取非相对论基组的精确基函数，仅优化收缩系数；文献 [62]~[68] 中的基组不仅优化收缩系数而且还优化基函数的指数。关于 DKH Hamilton 算符的更多讨论，读者可参见 Nakajima 和 Hirao 撰写的综述文献 [56]。

表 5.3 DKH Hamilton 算符的配套基组

Hamilton 算符	化学元素	基组特征	参考文献
DK2	1, 2, 11, 12	一般收缩	[57, 58]
	H, He, B~Ne, Al~Ar, Ga~Br	一般收缩, CC-PVXZ (X=D, T, Q, 5)	[59]
	1, 2, 3~18, La~Lu, Ac~Cm	一般收缩	[59]
	H~Xe	一般收缩, 有限核模型	[60]
	La~Lu, Hf~Hg	分段收缩	[59]
	H~Xe	分段收缩, DZP, TZP, QZP, 5ZP, 6ZP	[61]
	Ga~Kr, Sc~Zn, 11~12, Hg, Y~Pd, Hf~Pt	一般收缩，相关能一致	[62]
	La~Lu	分段收缩	[63]
	La~Lu, Ac~Lr	分段收缩	[64, 65]
DK3	H~Lr	一般收缩, 点电荷、有限核模型	[66, 67]
	H~Lr	分段收缩，有限核模型	[68]

注: DK2, DK3 分别指二次、三次 Douglas-Kroll 变换得到的 Hamilton 算符; 1, 2, $\cdots$, 18 表示 1 族、2 族、$\cdots$、18 族元素; 一般收缩和分段收缩见 6.10 节; 点电荷和有限核模型见 2.4 节。CC-PVXZ(X=D, T, Q, 5) 指相关能一致的价层极化基组 (分别为双 -Zeta、三 -Zeta、四 -Zeta 和五 -Zeta); DZP, TZP, QZP, 5ZP, 6ZP 指极化基组 (分别为双 -Zeta、价层三 -Zeta、价层四 -Zeta、价层五 -Zeta 和价层六 -Zeta)。

5.2.4 Wood-Boring Hamilton 算符

将单电子定态 Dirac 方程 (5.28) 式的第二式重写为

$$\Psi_{\mathrm{S}} = \frac{1}{2m_0c^2}\left(1+\frac{E-V}{2m_0c^2}\right)^{-1}(c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\Psi_{\mathrm{L}} \tag{5.129}$$

将其代回到方程组 (5.28) 式的第一式，得仅含大分量波函数的 Wood-Boring (WB) 方程

$$\left[(V-E)+\frac{1}{2m_0}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\cdot\frac{1}{1+\dfrac{E-V}{2m_0c^2}}\cdot(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\right]\Psi_{\mathrm{L}}=0 \tag{5.130}$$

Wood-Boring Hamilton 算符定义为

$$\hat{H}_{\mathrm{WB}}=V+\frac{1}{2m_0}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\cdot\frac{1}{1+\dfrac{E-V}{2m_0c^2}}\cdot(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) \tag{5.131}$$

下面考虑单电子原子，由氢原子的径向方程 (2.113) 式，得单电子原子的径向方程 (原子单位①)

$$\begin{aligned}\alpha(E-V)F&=\frac{\mathrm{d}G}{\mathrm{d}r}+\frac{k}{r}G\\ \frac{-2}{\alpha}\left[1+\frac{\alpha^2}{2}(E-V)\right]G&=\frac{\mathrm{d}F}{\mathrm{d}r}-\frac{k}{r}F\end{aligned} \tag{5.132}$$

式中，α 是精细结构常数；能量 E 中不包含静止能量。由方程组第二式，有

$$G=\frac{-\alpha}{2}B\left(\frac{\mathrm{d}F}{\mathrm{d}r}-\frac{k}{r}F\right) \tag{5.133}$$

其中

$$B=\frac{1}{1+\alpha^2/2(E-V)}$$

等式两边分别对 r 求导数，得

$$\begin{aligned}\frac{-\mathrm{d}G}{\mathrm{d}r}=\frac{\alpha}{2}\bigg[&B\frac{\mathrm{d}^2F}{\mathrm{d}r^2}+\left(-\frac{Bk}{r}+\frac{B^2\alpha^2}{2}\frac{\mathrm{d}V}{\mathrm{d}r}\right)\frac{\mathrm{d}F}{\mathrm{d}r}\\ &+\left(\frac{Bk}{r^2}-\frac{kB^2\alpha^2}{2r}\frac{\mathrm{d}V}{\mathrm{d}r}\right)F\bigg]\end{aligned} \tag{5.134}$$

把 (5.133) 式和 (5.134) 式的结果代入 (5.132) 式的第一个方程，得

$$\frac{1}{2}\frac{\mathrm{d}^2F}{\mathrm{d}r^2}+\frac{B\alpha^2}{4}\frac{\mathrm{d}V}{\mathrm{d}r}\frac{\mathrm{d}F}{\mathrm{d}r}+\left[(E-V)+\frac{\alpha^2}{2}(E-V)^2-\frac{l(l+1)}{2r^2}-\frac{B\alpha^2k}{4r}\frac{\mathrm{d}V}{\mathrm{d}r}\right]F=0 \tag{5.135}$$

将上式重组，得单电子原子的 WB 径向方程

$$\left(\hat{H}_{\mathrm{S}}+\hat{H}_{\mathrm{MV}}+\hat{H}_{\mathrm{D}}+\hat{H}_{\mathrm{SO}}\right)F=EF \tag{5.136}$$

式中，各项分别为

① 原子单位定义见附录C

$$\begin{aligned}\hat{H}_{\mathrm{S}} &= -\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{l(l+1)}{2r^2}+V\\ \hat{H}_{\mathrm{MV}} &= \frac{-\alpha^2}{2}(E-V)^2\\ \hat{H}_{\mathrm{D}} &= \frac{-B\alpha^2}{4}\frac{\mathrm{d}V}{\mathrm{d}r}\left(\frac{\mathrm{d}}{\mathrm{d}r}-\frac{1}{r}\right)\\ \hat{H}_{\mathrm{SO}} &= \frac{-B\alpha^2}{4}\frac{\mathrm{d}V}{\mathrm{d}r}\frac{1-k}{r}\end{aligned} \tag{5.137}$$

$\hat{H}_{\mathrm{S}}$ 为非相对论的径向方程，其余各项为相对论校正项。其中，$\hat{H}_{\mathrm{MV}}$ 称为质量–速度 (mass-velocity) 项，为电子动能项的高一级校正；$\hat{H}_D$ 称为 Darwin 项，是对势能项的修正；$\hat{H}_{\mathrm{SO}}$ 表示旋–轨耦合。如果 WB Hamilton 算符中不包括旋–轨耦合，所得 Hamilton 算符又称为 Cowan-Griffin (CG)Hamilton 算符。

将 $\pm k = 2\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}+1$(见 (2.86) 式) 代入 $\hat{H}_{\mathrm{SO}}$ 并和 (5.75) 式相比，我们发现 WB Hamilton 算符中的旋–轨耦合算符和 Pauli Hamilton 算符中的旋 - 轨耦合算符相差一个系数 B

$$\hat{H}_{\mathrm{SO}}^{\mathrm{WB}} = B\hat{H}_{\mathrm{SO}}^{\mathrm{Pauli}} \tag{5.138}$$

事实上，如果我们在 WB Hamilton 算符中利用 $\dfrac{1}{1+\dfrac{E-V}{2m_0c^2}}$ 的级数展开并只取展开式前两项，有

$$\begin{aligned}\frac{1}{1+\dfrac{E-V}{2m_0c^2}} &= 1-\frac{E-V}{2m_0c^2}+\cdots\\ &\approx 1-\frac{E-V}{2m_0c^2}\quad (E-V\ll 2m_0c^2)\end{aligned} \tag{5.139}$$

此时 WB 方程 (5.130) 式变为

$$\left[\frac{\hat{\boldsymbol{p}}^2}{2m_0}-\frac{\hat{\boldsymbol{p}}^2}{4m_0^2c^2}E+\frac{1}{4m_0^2c^2}\left(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\right)V\left(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\right)\right]\Psi_{\mathrm{L}} = (E-V)\,\Psi_{\mathrm{L}} \tag{5.140}$$

将上式左边第三项进行适当变形 (见 (5.35) 式) 也可得 Pauli Hamilton 算符 (见 (5.75) 式)。

为求解 (5.136) 式，Wood 和 Boring 提出如下步骤：

(1) 忽略旋–轨耦合及小分量 G 对 V 的影响，采用自洽场迭代的方法求解 F，并要求 F 是归一化的，即

$$\int F^2\mathrm{d}r = 1 \tag{5.141}$$

(2) 求得 F 后，再代入 (5.133) 式得到 G。

(3) 将旋–轨耦合作微扰处理。

他们发现，最后得到的 G 和真实值符合得相当好[69]。

需要指出的是，由于 WB Hamilton 算符中含有能量 E，求解 WB 方程得到的的轨道是不正交的，不利于基组展开求积分，所以 WB 方程一般用于原子的有限差分计算[70]，所得原子数据多用于拟合赝势和模型势 (见第 6 章)。

5.2.5　零级规则近似 Hamilton 算符

将 WB Hamilton 算符 (5.130) 式重写为

$$\left[(V-E)+(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\cdot\frac{c^2}{2m_0c^2-V}\cdot\frac{1}{1+\dfrac{E}{2m_0c^2-V}}\cdot(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\right]\Psi_{\mathrm{L}}=0 \tag{5.142}$$

考虑级数展开

$$\frac{1}{1+\dfrac{E}{2m_0c^2-V}}=1-\frac{E}{2m_0c^2-V}+\cdots \tag{5.143}$$

当 $\dfrac{E}{2m_0c^2-V}\ll 1$ 时，上式可近似为 1，相当于只取级数展开中 $n=0$ 时的展开项，所以该近似又称为零级规则近似 (zeroth-order regular approximation，ZORA)。于是 (5.142) 式变为

$$\left[(V-E)+(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\cdot\frac{c^2}{2m_0c^2-V}\cdot(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\right]\Psi_{\mathrm{L}}=0 \tag{5.144}$$

下面归一化上式中的大分量波函数 Ψ_{L}。

设归一化算符 $\hat{O}$，将 $\hat{O}$ 作用于 Ψ_{L}，有

$$\int \Psi_{\mathrm{L}}^{\dagger}\hat{O}^{\dagger}\hat{O}\,\Psi_{\mathrm{L}}\mathrm{d}^3\tau=1 \tag{5.145}$$

由 Dirac 方程的波函数归一化条件

$$\int(\Psi_{\mathrm{L}}^{\dagger}\Psi_{\mathrm{L}}+\Psi_{\mathrm{S}}^{\dagger}\Psi_{\mathrm{S}})\mathrm{d}^3\tau=1 \tag{5.146}$$

将 (5.129) 式代入上式，得

$$\int \Psi_{\mathrm{L}}^{\dagger}(1+\hat{W}^{\dagger}\hat{W})\Psi_{\mathrm{L}}\mathrm{d}^3\tau=1 \tag{5.147}$$

其中 $\hat{W}$ 定义为

$$\hat{W}=\frac{c}{2m_0c^2-V}\cdot\frac{1}{1+\dfrac{E}{2m_0c^2-V}}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) \tag{5.148}$$

比较 (5.147) 式和 (5.145) 式，求得归一化算符 $\hat{O}$

$$\hat{O} = \sqrt{1 + \hat{W}^{\dagger}\hat{W}} \tag{5.149}$$

将 $\hat{O}$ 作级数展开并联合 (5.143) 式，得

$$\begin{aligned}
\hat{O} &= 1 + \frac{1}{2}\hat{W}^{\dagger}\hat{W} + \cdots \\
&= 1 + \frac{1}{2}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\frac{c^2}{(2m_0c^2 - V)^2}\left(1 + \frac{E}{2m_0c^2 - V}\right)^{-2}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) + \cdots \\
&= 1 + \frac{1}{2}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c^2}{2m_0c^2 - V}\right)\frac{1}{2m_0c^2 - V}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) \\
&\quad -(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c^2}{2m_0c^2 - V}\right)\left(\frac{1}{2m_0c^2 - V}\right)\left(\frac{E}{2m_0c^2 - V}\right)(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) + \cdots
\end{aligned} \tag{5.150}$$

如果只考虑零级展开，上式约为 1。也就是说方程 (5.144) 式中的 Ψ_{L} 在 ZORA 下是归一化的。我们将方程 (5.144) 式又称为 ZORA 方程。ZORA Hamilton 算符可写为

$$\hat{H}_{\mathrm{ZORA}} = V + (\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,\frac{c^2}{2m_0c^2 - V}\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) \tag{5.151}$$

ZORA 由 Chang, Pélissier, Durand[71] 于 1986 年首先提出，文献中 ZORA Hamilton 算符也称为 CPD Hamilton 算符。

将 $(2m_0c^2 - V)^{-1} = \dfrac{1}{2m_0c^2}\left(1 - \dfrac{V}{2m_0c^2}\right)^{-1}$ 作级数展开，再代入上式得

$$\begin{aligned}
\hat{H}_{\mathrm{ZORA}} &= V + \frac{1}{2m_0}\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left[1 + \frac{V}{2m_0c^2} + \frac{V^2}{4m_0^2c^4} + \cdots\right](\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}) \\
&= V + \frac{\hat{\boldsymbol{p}}^2}{2m_0} + \frac{(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\,V\,(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})}{4m_0^2c^2} + \cdots
\end{aligned} \tag{5.152}$$

与导出 Pauli Hamilton 算符的 (5.140) 式相比，我们发现 ZORA Hamilton 算符缺少一项，即 $\dfrac{-\hat{\boldsymbol{p}}^2E}{4m_0^2c^2}$。此项显然大于 0，所以求解 ZORA Hamilton 算符得到的能量一般偏低。也因为不包含此项， ZORA Hamilton 算符不存在奇点，而包含该项的 Pauli Hamilton 算符却存在奇点。另外，该项给出了电子动能的高级校正 (质量–速度项) 及势能的部分校正 (Darwin 项)，所以缺少此项的 ZORA Hamilton 算符虽然可以得到正确的旋–轨耦合，但不足以描述不包括自旋的相对论校正。

下面我们将 ZORA Hamilton 算符与 Dirac Hamilton 算符进行比较。取 V 为原子的 Coulomb 势，大分量波函数的 Dirac 方程可写为

$$\left[\frac{-Z}{r} + (\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c^2}{2m_0c^2 + E_{\mathrm{D}} + Z/r}\right)(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\right]\Psi_{\mathrm{L}}(\boldsymbol{r}) = E_{\mathrm{D}}\,\Psi_{\mathrm{L}}(\boldsymbol{r}) \tag{5.153}$$

式中，Z 为核电荷数。定义变量 $\boldsymbol{r}' = \lambda \boldsymbol{r}$，其中

$$\lambda = \frac{2m_0c^2 + E_{\mathrm{D}}}{2m_0c^2} \tag{5.154}$$

则有，$\hat{\boldsymbol{p}} = \lambda \hat{\boldsymbol{p}}'$，将 (5.153) 式重写为

$$\left[\frac{-Z}{r'} + (\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{p}}')\left(\frac{c^2}{2m_0c^2 + Z/r'}\right)(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{p}}')\right] \Psi_{\mathrm{L}}(\boldsymbol{r}'/\lambda) = \frac{2m_0c^2 E_{\mathrm{D}}}{2m_0c^2 + E_{\mathrm{D}}} \Psi_{\mathrm{L}}(\boldsymbol{r}'/\lambda) \tag{5.155}$$

与 ZORA 方程 (5.144) 式相比，我们发现 ZORA 波函数即为新坐标系下 Dirac 方程的大分量波函数，ZORA 能量 E_{ZORA} 与 Dirac 能量 E_{D} 之间的关系为

$$E_{\mathrm{ZORA}} = \frac{2m_0c^2 E_{\mathrm{D}}}{2m_0c^2 + E_{\mathrm{D}}} = E_{\mathrm{D}}\left[1 + \frac{E_{\mathrm{D}}}{2m_0c^2}\right]^{-1} = 2m_0c^2\left[1 + \frac{2m_0c^2}{E_{\mathrm{D}}}\right]^{-1} \tag{5.156}$$

观察上式我们发现

(1) 当 $E_{\mathrm{D}} \to +\infty$ 时， $E_{\mathrm{ZORA}} \to 2m_0c^2$ 。Dirac 方程的正能区间由 $(0,\ \infty)$ 压缩至 $(0,\ 2m_0c^2)$。

(2) 对束缚态解，即当 $-m_0c^2 < E_{\mathrm{D}} < 0$ 时，$-2m_0c^2 < E_{\mathrm{ZORA}} < 0$。

(3) 当 $E_{\mathrm{D}} < -2m_0c^2$ 时，$E_{\mathrm{ZORA}} > 0$；当 $E_{\mathrm{D}} \to -\infty$ 时，$E_{\mathrm{ZORA}} \to 2m_0c^2$；当 $E_{\mathrm{D}} \to -2m_0c^2$ 时，$E_{\mathrm{ZORA}} \to +\infty$。由此可见，ZORA 方程将 Dirac 方程的负能解 $(-\infty, -2m_0c^2)$ 映射到正能区域 $(2m_0c^2, \infty)$。

(4) 当 $E_{\mathrm{D}} = -2m_0c^2$ 时，方程存在奇点，此时 ZORA 波函数为 0。

由于以上特性，ZORA Hamilton 算符具有变分稳定性，不会出现变分塌陷。注意该结论仅当 $Z < c$ 时成立，原因是采用点电荷模型的 Dirac 方程在 $Z > c$ 时不存在束缚态解 (见 2.3.3 节)，详细证明请见 van Lenthe 的博士论文①。

1. ZORA Hamilton 算符的求解

由于势能 V 出现在分母中，实际计算中我们需要求解势能相关项的逆，一种解决方案是采用数值积分技术，这时的势能 V 是网格上的数值点，也就不存在求逆的困难。这个方法很适合 DFT 理论，因为在 DFT 计算中交换–相关泛函的值就是网格上的数值点，Coulomb 势也很容易数值化。另外一种办法是利用 RI (resolution of identity) 近似求得 Hamilton 算符矩阵表达，再求表达矩阵的逆[72]。在该方法中由 RI 近似[73] 引入的辅助基组 φ 与大分量基组 ϕ (用来展开 ZORA 方程的波函数 Ψ_{L}) 的关系为

$$\varphi = (\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{p}})\phi \frac{\mathrm{M}[\hat{T}]^{-1/2}}{(2m_0)^{1/2}} \tag{5.157}$$

① van Lenthe E. The ZORA Equation. Amsterdam: Vrije Universiteit, 1996

M$[\hat{T}]$ 为动能算符 $\left(\hat{T}=\dfrac{(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})}{2m_0}\right)$ 的矩阵表示。与基函数动能平衡条件 (见 (5.8) 式) 比较可知，此辅助基组即为小分量基组。也就是说虽然 ZORA Hamilton 算符省略了小分量波函数，但仍需要小分量基函数。

由于势能出现在分母中，所以 ZORA 方程的解和所选取的势能零点 (或称规范) 相关。以类氢原子为例，将势能零点平移一个常量 Δ

$$V^{\Delta}=-\frac{Z}{r}+\Delta \tag{5.158}$$

对 Dirac 方程而言，势能零点平移后能量本征值 E_{D} 也平移常量 Δ，而 ZORA 方程的能量在势能零点平移后变为[74]

$$E_{\mathrm{ZORA}}^{\Delta}=E_{\mathrm{ZORA}}+\Delta-\frac{E_{\mathrm{ZORA}}\Delta}{2m_0c^2} \tag{5.159}$$

当 $|E_{\mathrm{ZORA}}|\ll 2m_0c^2$ 时 (如价电子轨道)，式中的第三项很小可以忽略。我们通常选取无穷远处为势能零点，读者也许因此认为 ZORA 方程对势能零点的依赖不会对计算结果产生影响，其实不然。表 5.4 列出了 van Lenthe 等计算得到的金的电离能[74]，其中 ZORA 方程的结果和 Dirac 方程的结果相差甚远，见表 5.4 中的第二列和第三列，这是由于外层电子电离对壳电子来说相当于势能平移了一个常数，影响了电离能的计算。

表 5.4 van Lenthe 等计算得到的金的电离能 (Hartree)[75]

电离态	全电子计算		价电子计算	
	Dirac	ZORA	Dirac	ZORA
$6s_{1/2}^1\to 6s_{1/2}^0$	0.3286	0.1650	0.3285	0.3279
$5d_{5/2}^6\to 5d_{5/2}^5$	0.3992	0.1745	0.3999	0.3991
$6s_{1/2}^1 5d_{5/2}^6\to 6s_{1/2}^0 5d_{5/2}^5$	1.0609	0.6443	1.0618	1.0603

注：以上计算均采用了 $X\alpha$ DFT 方法[76]； 1 Hartree = 4.3597×10^{-18} J, 下同。

考虑到势能平移仅对内层电子影响较大 (见 (5.159) 式)，在计算中可冻结内层轨道，只对价电子作计算。此时电子密度分为两部分，即内层电子密度和价电子密度；内层电子的电子密度可选取全电子 Dirac 计算的结果。从表 5.4 可以看出，采用内层电子冻结近似后 ZORA 的计算结果和 Dirac 方程的结果十分接近，误差小于 0.001 Hartree[74]。必须指出的是，虽然在 WB Hamilton 算符中势能也出现在分母中 (见 (5.131) 式)，但 WB Hamilton 算符中含有能量 E，在自洽场计算中会补偿势能零点的变化而带来的影响。

2. ZORA Hamilton 算符的校正

如果取级数展开式 (5.143) 式和 (5.150) 式中的前两项，则得到一级规则近似

(first-order regular approximation，FORA) Hamilton 算符[50]

$$\begin{aligned}\hat{H}_{\text{FORA}} &= \hat{H}_{\text{ZORA}} + \hat{H}' \\ &\equiv \hat{H}_{\text{ZORA}} + \frac{1}{2}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c}{2m_0c^2-V}\right)^2(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\hat{H}_{\text{ZORA}} \\ &\quad -\frac{1}{2}\hat{H}_{\text{ZORA}}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c}{2m_0c^2-V}\right)^2(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\end{aligned} \tag{5.160}$$

由于 $\hat{H}'$ 中包含算符 $\hat{\boldsymbol{p}}^4$，所以 FORA Hamilton 算符和 Pauli Hamilton 算符一样，一般不用于变分计算，$\hat{H}'$ 可以作微扰处理，一级微扰能量可写为

$$\begin{aligned}E_1 &= \langle \Psi_{\text{ZORA}}|\hat{H}'|\Psi_{\text{ZORA}}\rangle \\ &= -\frac{1}{2}\langle \Psi_{\text{ZORA}}|\left[(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c}{2m_0c^2-V}\right)^2(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\hat{H}_{\text{ZORA}}\right. \\ &\quad \left.+\hat{H}_{\text{ZORA}}(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c}{2m_0c^2-V}\right)^2(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\right]|\Psi_{\text{ZORA}}\rangle \\ &= -E_{\text{ZORA}}\langle \Psi_{\text{ZORA}}|(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\left(\frac{c}{2m_0c^2-V}\right)^2(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})|\Psi_{\text{ZORA}}\rangle \\ &= -E_{\text{ZORA}}\frac{\langle \Psi_{\text{S}}|\Psi_{\text{S}}\rangle}{\langle \Psi_{\text{L}}|\Psi_{\text{L}}\rangle}\end{aligned} \tag{5.161}$$

由小分量波函数与大分量波函数的关系式，即 (5.129) 式，可估计出零级展开时二者的模之比约为 $\dfrac{\hat{\boldsymbol{p}}^2}{2m_0c^2}$，代入上式得一级微扰能量为

$$E_1 \approx -\frac{E_{\text{ZORA}}\hat{\boldsymbol{p}}^2}{4m_0^2c^2} \tag{5.162}$$

在前面的讨论中，我们已经发现 ZORA Hamilton 算符比 Pauli Hamilton 算符少一项 $\dfrac{-E\hat{\boldsymbol{p}}^2}{4m_0^2c^2}$。由上式知，ZORA 能量的一级校正即为所缺项对能量的贡献。

一级规则展开近似还可以继续扩展至二级、三级以至无穷，有兴趣的读者可参考相关文献 [75]。

5.2.6 精确的二分量方法

以上我们讲到的二分量方法，均是采用直接忽略小分量或酉变换的方法将组态空间的四分量 Dirac 算符变换为二分量算符，实际应用时再引入基组将其离散化得到可解的矩阵方程。另一种二分量方法是先引入基组，得到 Dirac 方程的矩阵表示，再通过变换消除小分量得到二分量的矩阵方程，这就是所谓的“精确的二分量 (exact two-component，X2C) 方法”，又称“精确的准相对论矩阵理论 (exact matrix quasi-relativistic theory，XQR)”[77−81]。

假设单电子在势阱 V 中运动，电子的定态 Dirac 方程可写为 (原子单位)

$$\begin{aligned} c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\mathrm{S}}+[V-E]\,\Psi_{\mathrm{L}}&=0\\ c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\mathrm{L}}+\left[V-2m_0c^2-E\right]\Psi_{\mathrm{S}}&=0 \end{aligned} \tag{5.163}$$

式中，电子能量 E 为相对于静止能量 (m_0c^2) 的能量。
设

$$\Psi_{\mathrm{S}}=\hat{X}\Psi_{\mathrm{L}} \tag{5.164}$$

由方程组 (5.163) 式中的第二式，得

$$\begin{aligned} E\Psi_{\mathrm{S}}&=c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\mathrm{L}}+\left[V-2m_0c^2\right]\Psi_{\mathrm{S}}\\ E\hat{X}\Psi_{\mathrm{L}}&=c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\mathrm{L}}+\left[V-2m_0c^2\right]\hat{X}\Psi_{\mathrm{L}} \end{aligned} \tag{5.165}$$

将方程组 (5.163) 式中的第一式左乘 $\hat{X}$，并考虑到上式和 (5.164) 式，得

$$\hat{X}V\Psi_{\mathrm{L}}+\hat{X}c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\hat{X}\Psi_{\mathrm{L}}=c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\mathrm{L}}+\left[V-2m_0c^2\right]\hat{X}\Psi_{\mathrm{L}} \tag{5.166}$$

将上式整理，得

$$\left[\hat{X},V\right]\Psi_{\mathrm{L}}+\hat{X}c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\hat{X}\Psi_{\mathrm{L}}=c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\Psi_{\mathrm{L}}-2m_0c^2\hat{X}\Psi_{\mathrm{L}} \tag{5.167}$$

对所有可能的 Ψ_{L}，$\hat{X}$ 均需满足上式，所以 $\hat{X}$ 必满足以下方程[82]

$$2m_0c^2\hat{X}=c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}-\left[\hat{X},V\right]-c\hat{X}\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\hat{X} \tag{5.168}$$

这是一个关于 $\hat{X}$ 的二次方程，有两个独立的解，分别对应于电子和正电子[55]。为保证我们得到电子解，还需寻找 $\hat{X}$ 的约束条件。
由方程组 (5.163) 式的第二式，可得

$$\hat{X}=\frac{\Psi_{\mathrm{S}}}{\Psi_{\mathrm{L}}}=\frac{c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}}{E-V+2m_0c^2} \tag{5.169}$$

在非相对论极限条件下，$\hat{X}\rightarrow\dfrac{\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}}{2m_0c}$，此即为对 $\hat{X}$ 的约束条件。将 (5.164) 式代入方程组 (5.163) 式中的第一式，得

$$\hat{L}=V+c\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\hat{X},\quad \hat{L}\Psi_{\mathrm{L}}=E\Psi_{\mathrm{L}} \tag{5.170}$$

式中，$\hat{L}$ 为非厄米的准相对论算符。这样，我们就将 Dirac 方程的求解问题转化为对上式及 (5.168) 式的求解问题。然而，(5.168) 式的求解并不比求解 Dirac 方程简单，仅对一些特殊形式的势能 V，我们能得到封闭形式的解 $\hat{X}$ [55]。

在 5.1.1 节我们已经提到了 Dirac 方程的矩阵表达式 (见 (5.2) 式)，由 $\hat{X}$ 的约束条件 (5.169) 式知用以展开 Ψ_L 和 Ψ_S 的基函数 $\{\phi_i\}, \{\chi_j\}$ 必须满足基组动能平衡条件 (5.8) 式。此时可将 (5.164) 式表示为

$$\Psi_{\text{S},k} = \hat{\tilde{X}}\,\Psi_{\text{L},k} \quad (k=1,2,\cdots,m) \tag{5.171}$$

式中，$\hat{\tilde{X}}$ 对每一个 $\Psi_{\text{L},k}$ 都相同。由于精确的波函数并不能用通用基组展开，所以式中的 $\hat{\tilde{X}}$ 并不等于 (5.164) 式中的 $\hat{X}$，可看成是对 $\hat{X}$ 的近似。

接下来将 Ψ_L 和 Ψ_S 用基函数 $\{\phi_i\}, \{\chi_j\}$ 展开

$$\begin{aligned} \Psi_\text{L} &= \sum_i a_i\phi_i \\ \Psi_\text{S} &= \sum_i b_i\chi_i = \sum_i b_i(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}})\phi_i = \hat{\tilde{X}}\,\Psi_\text{L} = \sum_i a_i\hat{\tilde{X}}\phi_i \end{aligned} \tag{5.172}$$

于是我们可以得到 Dirac 方程在动能平衡基下的矩阵表示[78]

$$\begin{pmatrix} \text{V} & 2m_0c\text{W} \\ 2m_0c\text{W} & \text{U}-4m_0^2c^2\text{W} \end{pmatrix}\begin{pmatrix}\text{a}\\ \text{b}\end{pmatrix} = E\begin{pmatrix}\text{S} & 0 \\ 0 & 2m_0\text{W}\end{pmatrix}\begin{pmatrix}\text{a}\\ \text{b}\end{pmatrix} \tag{5.173}$$

式中，相关矩阵元为

$$\begin{aligned} V_{ij} = \langle\phi_i|V|\phi_j\rangle, &\quad U_{ij} = \langle\phi_i|\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}V\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}|\phi_j\rangle \\ S_{ij} = \langle\phi_i|\phi_j\rangle, &\quad W_{ij} = \langle\phi_i|\frac{\hat{\boldsymbol{p}}^2}{2m_0}|\phi_j\rangle \end{aligned} \tag{5.174}$$

定义 $m\times m$ 维矩阵 X，其矩阵元为

$$X_{ij} = \langle\phi_i|\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{p}}\hat{\tilde{X}}|\phi_j\rangle \tag{5.175}$$

矩阵 X 和 W, a, b 存在如下关系

$$\text{Wb} = \frac{1}{2m_0}\text{Xa} \ \rightarrow\ \text{b} = \frac{1}{2m_0}\text{W}^{-1}\text{Xa} \tag{5.176}$$

利用上式，我们可以消除 Dirac 矩阵方程 (5.173) 式中的 b，得非厄米的准相对论 Hamilton 矩阵方程

$$\text{L} = \text{V} + c\text{X}, \quad \text{La} = E\text{Sa} \tag{5.177}$$

上式也可转换为厄米的准相对论 Hamilton 矩阵方程，感兴趣的读者可阅读刘文剑和 Kutzelnigg 的文章[83]。

(5.177) 式中 X 满足以下矩阵方程

$$\text{X} = c^{-1}\text{W} + \frac{1}{4m_0^2c^2}\text{UW}^{-1}\text{X} - \frac{1}{2m_0c^2}\text{XS}^{-1}\text{L} \tag{5.178}$$

如果忽略上式中量级在 $O(c^{-2})$ 的小项，则可得到 X 和 L 的非相对论极限 (nrel) 表达式

$$\mathrm{X}_{\mathrm{nrel}} = c^{-1}\mathrm{W}, \quad \mathrm{L}_{\mathrm{nrel}} = \mathrm{V} + \mathrm{W} \tag{5.179}$$

(5.178) 式是一个关于 X 的非均匀的二次方程，只能用线性 (或二次) 迭代的方法进行求解[83]。

本节介绍的 X2C 方法由刘文剑和 Kutzelnigg 于 2005 年首先发表。目前，Jensen 等已将 X2C 方法在量子化学程序 DIRAC 11 [14] 中实现，进入到推广使用阶段。X2C 方法与其他二分量 Hamilton 算符，如 Pauli Hamilton 算符、DKH Hamilton 算符、ZORA Hamilton 算符等的详细讨论读者可参阅刘文剑等的综述文章[83, 84]。

参 考 文 献

[1] Dyall K G, Grant I P, Johnson C T, et al. GRASP: a general-purpose relativistic atomic structure program. Comput Phys Commun, 1989, 55: 425～456.

[2] Desclaux J P, Indelicato P. MCDFGME: a Multiconfiguration Dirac Fock and General Matrix Elements Program. http://dirac.spectro.jussieu.fr/mcdf [2013-04-25].

[3] Vilkas M J, Koc K, Ishikawa Y. Second-order multiconfigurational Dirac-Fock self-consistent field and multireference configuration interaction calculations on beryllium and beryllium-like Ne^{6+}. Chem Phys Lett, 1997, 280: 167～176.

[4] Vilkas M J, Ishikawa Y, Koc K. Quadratically convergent multiconfiguration Dirac-Fock and multireference relativistic configuration-interaction calculations for many-electron systems. Phys Rev E, 1998, 58: 5096～5110.

[5] Dzuba V A, Flambaum V V, Kozlov M G. Combination of the many-body perturbation theory with the configuration-interaction method. Phys Rev A, 1996, 54: 3948～3959.

[6] Vilkas M J, Ishikawa Y, Koc K. Relativistic multireference many-body perturbation theory for quasidegenerate systems: Energy levels of ions of the oxygen isoelectronic sequence. Phys Rev A, 1999, 60: 2808～2821.

[7] Dzuba V A, Flambaum V V, Sushkov O P. Relativistic many-body calculations of energy levels and of fine-structure intervals in the caesium atom. J Phys B, 1983, 16: 715～722.

[8] Blundell S A, Johnson W R, Sapirstein J. 3rd-order many-body perturbation-theory calculations of the ground-state energies of cesium and thalium. Phys Rev A, 1990, 42: 3751～3762.

[9] Salomonson S, Lindgren I, Mårtensson A M. Numerical many-body perturbation calculations on Be-like systems using a multi-configurational model space. Phys Scr, 1980, 21: 351～356.

[10] Blundell S A, Johnson W R, Sapirstein J. Relativistic all-order calculations of energies and matrix elements in cesium. Phys Rev A, 1991, 43: 3407~3418.

[11] Salomonson S, Öster P. Relativistic all-order pair functions from a discretized single-particle Dirac Hamiltonian. Phys Rev A, 1989, 40: 5548~5558.

[12] Eliav E, Kaldor U, Ishikawa Y. Open-shell relativistic coupled-cluster method with Dirac-Fock-Breit wave functions: Energies of the gold atom and its cation. Phys Rev A, 1994, 49: 1724~1729.

[13] Das B P, Latha K V P, Sahoo B K, et al. Relativistic and correlation effects in atoms. J Theor Comp Chem, 2005, 1: 1~20.

[14] Jensen H J, Saue T, Visscher L, et al. DIRAC: a relativistic *ab initio* electronic structure program. http://wiki.chem.vu.nl/dirac/index.php/Dirac_Program [2013-04-25].

[15] (a) Visscher L, Visser O, Aerts P J C, et al. Relativistic quantum chemistry: the MOLFDIR program package. Comp Phys Commun, 1994, 81: 120~144.

(b) Visscher L, de Jong W A, Merenga H, et al. Relativistic quantum chemistry: the MOLFDIR program package. // Clementi E, Corongiu G. Methods and Techniques for Computational Chemistry, METECC-5. Strasbourg: Universite L Pasteur, 1995: 169~218.

(c) Pernpointner M, de Jong W A, Visscher L, et al. The paralleization of four-component calculations: I. Integral generation, SCF and four-index transformation in the Dirac-Fock package MOLFDIR. J Comput Chem, 2000, 21: 1176~1186.

[16] Yanai T, Nakano H, Nakajima T, et al. UTCHEM—a program for *ab initio* quantum chemistry.//Sloot P M A, Abramson D, Bogdanov A V, et al. Lecture Notes in Computer Science. Computer Science–ICCS 2003. Vol 2660. Berlin: Springer-Verlag, 2003: 84~95.

[17] Matsuoka O, Watanabe Y. PROPHET4R: Four-component relativistic atomic and molecular program suite. // Hirao K, Ishikawa Y. Advances in Relativistic Molecular Theory. Recent Advances in Computational Chemistry. Vol 5. Singapore: World Scientific Pub Co Inc, 2004: 247~256.

[18] Quiney H M, Skaane H, Grant I P. *Ab initio* relativistic quantum chemistry: four-components good, two-components bad! Adv Quant Chem, 1999, 32: 1~49.

[19] Liu W J, Hong G Y, Dai D D, et al. The Beijing four-component density functional program package (BDF) and its application to EuO, EuS, YbO and YbS. Theor Chem Acc, 1997, 96: 75~83.

[20] Drake G W F, Goldman S P. Application of discrete-basis-set methods to the Dirac equation. Phys Rev A, 1981, 23: 2093~2098.

[21] Talman J D. Minimax principle for the Dirac equation. Phys Rev Lett, 1986, 57: 1091~1094.

[22] Hylleraas E A, Undheim B. Numerische Berechnung der ^{2}S-Terme von ortho- and para-Helium. Z Phys, 1930, 65: 759~772.

[23] MacDonald J K L. Successive approximations by the Rayleigh-Ritz variation method. Phys Rev, 1933, 43: 830~833.

[24] Schwarz W H E, Wallmeier H. Basis set expansions of relativistic molecular wave equations. Mol Phys, 1982, 46: 1045~1061.

[25] Brown G E, Ravenhall D G. On the interaction of two electrons. Proc Roy Soc London A, 1951, 208: 552~559.

[26] Kutzelnigg W. The relativistic many body problem in molecular theory. Phys Scr, 1987, 36: 416~431.

[27] Sucher J. Foundations of the relativistic theory of many-electron atoms. Phys Rev A, 1980, 22: 348~362.

[28] Furry W H. On bound states and scattering in positron theory. Phys Rev, 1951, 81: 115~124.

[29] Mittleman M H. Theory of relativistic effects on atoms: Configuration-space Hamiltonian. Phys Rev A, 1981, 24: 1167~1175.

[30] Jensen H J, Dyall K G, Saue T, et al. Relativistic four-component multiconfigurational self-consistent-field theory for molecules: Formalism. J Chem Phys, 1996, 104: 4083~4098.

[31] Sapirstein J. Theoretical methods for the relativistic atomic many-body problem. Rev Mod Phys, 1998, 70: 55~76.

[32] Kaldor U. The Fock space coupled cluster method: theory and application. Theor Chim Acta, 1991, 80: 427~439.

[33] Kaldor U, Eliav E. High-accuracy calculations for heavy and super-heavy elements. Adv Quantum Chem, 1998, 31: 313~336.

[34] Infante I, Eliav E, Vilkas M J, et al. A Fock space coupled cluster study on the electronic structure of the UO_2, UO_2^+, U^{4+}, and U^{5+} species. J Chem Phys, 2007, 127: 124308-1~124308-12.

[35] Saue T, Visscher L. Four-component electronic structure methods for molecules. // Kaldor U, Wilson S. Theoretical Chemistry and Physics of Heavy and Super-Heavy Elements. Dordrecht: Springer Science+Business Media B V, 2003: 211~262.

[36] Hohenberg P, Kohn W. Inhomogeneous electron gas. Phys Rev, 1964, 136: B864~B871.

[37] Kohn W, Sham L J. Self-consistent equations including exchange and correlation effects. Phys Rev, 1965, 140: A1133~A1138.

[38] Vosko S H, Wilk L, Nusair M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. Can J Phys, 1980, 58: 1200~1211.

[39] Perdew J P, Zunger A. Self-interaction correction to density-functional approximations for many-electron systems. Phys Rev B, 1981, 23: 5048~5079.

[40] Ceperley D M, Alder B J. Ground state of the electron gas by a stochastic method. Phys Rev Lett, 1980, 45: 566~569.

[41] Jones R D, Gunnarsson O. The density functional formalism, its applications and prospects. Rev Mod Phys, 1989, 61: 689~746.

[42] Ma S H, Brueckner K A. Correlation energy of an electron gas with a slowly varying high density. Phys Rev, 1965, 165: 18~31.

[43] Becke A D. Exploring the limits of gradient corrections in density functional theory. J Comput Chem, 1999, 20: 63~69.

[44] Vignale G, Rasolt M. Density-functional theory in strong magnetic fields. Phys Rev Lett, 1987, 59: 2360~2363.

[45] Vignale G, Rasolt M. Current- and spin-density-functional theory for inhomogeneous electronic systems in strong magnetic fields. Phys Rev B, 1988, 37: 10685~10696.

[46] Engel E. Relativistic density functional theory: Foundations and basic formalism. // Schwerdtfeger P. Relativistic Electronic Structure Theory, Part1: Fundamentals. Theoretical and Computational Chemistry. Vol 11. Amsterdam: Elsevier Science B V, 2002: 523~613.

[47] MacDonald A H, Vosko S H. A relativistic density functional formalism. J Phys C, 1979, 12: 2977~2990.

[48] Engel E, Keller S, Bonetti A F, et al. Local and nonlocal relativistic exchange-correlation energy functionals: Comparison to relativistic optimized-potential-model results. Phys Rev A, 1995, 52: 2750~2764.

[49] Engel E, Keller S, Dreizler R M. Generalized gradient approximation for the relativistic exchange-only energy functional. Phys Rev A, 1996, 53: 1367~1374.

[50] van Lenthe E, Baerends E J, Snijders J G. Relativistic regular two-component Hamiltonians. J Chem Phys, 1993, 99: 4597~4610.

[51] Tarczay G, Császár A G, Klopper W, et al. Anatomy of relativistic energy corrections in light molecular systems. Mol Phys, 2001, 99: 1769~1794.

[52] Douglas M, Kroll N M. Quantum electrodynamical corrections to the fine structure of Helium. Ann Phys, 1974, 82: 89~155.

[53] Hess B A. Relativistic effects in heavy-element chemistry. Ber Bunsenges Phys Chem, 1997, 101: 1~10.

[54] Hess B A, Marian C M, Wahlgren U, et al. A mean-field spin-orbit method applicable to correlated wavefunctions. Chem Phys Lett, 1996, 251: 365~371.

[55] Wolf A, Reiher M, Hess B. Two-component methods and the generalised Douglas-Kroll transformation. // Schwerdtfeger P. Relativistic Eleltronic Structure Theory, Part I: Fundamentals, Theoretical and Computational Chemistry. Vol 11. Amsterdam:

Elsevier Science B V, 2002: 622～659.

[56] Nakajima T, Hirao K. The Douglas-Kroll-Hess approach. Chem Rev, 2012, 112: 385～402.

[57] Miadoková I, Kello V, Sadlej A J. Standardized basis sets for high-level-correlated relativistic calculations of atomic and molecular electric properties in the spin-averaged Douglas-Kroll approximation 2 Groups Ⅰa and Ⅱa. Theor Chem Acc, 1997, 96: 166～175.

[58] Kello V, Sadlej A J. Standardized basis sets for high-level-correlated relativistic calculations of atomic and molecular electric properties in the spin-averaged Douglas-Kroll (no-pair) approximation 1 Groups Ⅰb and Ⅱb. Theor Chim Acta, 1996, 94: 93～104.

[59] Schuchardt K L, Didier B T, Elsethagen T, et al. Basis set exchange v1.2.2. https://bse.pnl.gov/bse/portal[2012-06-19].

[60] Haiduke R L A, Comar M Jr, da Silva A B F. The employment of relativistic adapted Gaussian basis sets in Douglas-Kroll-Hess scalar calculations with diatomic molecules. Chem Phys, 2006, 331: 173～177.

[61] Jorge F E. Gaussian basis sets for Douglas-Kroll-Hess calculations. http://www.cce.ufes. br/qcgv/pub[2012-06-19].

[62] Peterson K A. Correlation-consistent basis sets. http://tyr0.chem.wsu.edu/～kipeters/basis.html[2012-06-20].

[63] Dolg M. Segmented contracted Douglas-Kroll-Hess adapted basis sets for lanthanides. J Chem Theor Comput, 2011, 7: 3131～3142.

[64] Pantazis D A, Neese F. All-electron scalar relativistic basis sets for the lanthanides. J Chem Theor Comput, 2009, 5: 2229～2238.

[65] Pantazis D A, Neese F. All-electron scalar relativistic basis sets for the actinides. J Chem Theor Comput, 2011, 7: 677～684.

[66] Tsuchiya T, Abe M, Nakajima T, et al. Accurate relativistic Gaussian basis sets for H through Lr determined by atomic self-consistent field calculations with the third-order Douglas-Kroll approximation. J Chem Phys, 2001, 115: 4463～4472.

[67] Nakajima T, Hirao K. Accurate relativistic Gaussian basis sets determined by the third-order Douglas-Kroll approximation with a finite-nucleus model. J Chem Phys, 2002, 116: 8270～8275.

[68] Takeshi N T, Sekiya M, Osanai Y. Segmented Gaussian basis sets. http://setani.sci.hokudai. ac.jp/sapporo[2012-06-20].

[69] Wood J H, Boring A M. Improved Pauli Hamiltonian for local-potential problems. Phys Rev B, 1978, 18: 2701～2711.

[70] Froese-Fischer C. The Hartree-Fock Method for Atoms. New York: John Wiley & Sons Inc, 1977.

[71] Chang C H, Pélissier M, Durand P H. Regular two-component Pauli-like effective

Hamiltonian in Dirac theory. Phys Scr, 1986, 34: 394~404.

[72] Filatov M. On representation of the Hamiltonian matrix elements in relativistic regular approximation. Chem Phys Lett, 2002, 365: 222~231.

[73] Eichkorn K, Treutler O, Ohm H, et al. Auxiliary basis-sets to approximate Coulomb potentials. Chem Phys Lett, 1995, 240: 283~290.

[74] van Lenthe E, Baerends E J, Snijders J G. Relativistic total energy using regular approximations. J Chem Phys, 1994, 101: 9783~9792.

[75] Dyall K G, van Lenthe E. Relativistic regular approximations revisited: An infinite-order relativistic approximation. J Chem Phys, 1999, 111: 1366~1372.

[76] Slater J C. A simplification of the Hartree-Fock method. Phys Rev, 1951, 81: 385~390.

[77] Dyall K G. Interfacing relativistic and nonrelativistic methods I. Normalized elimination of the small component in the modified Dirac equation. J Chem Phys, 1997, 106: 9618-1~9618-9.

[78] Kutzelnigg W, Liu W. Quasirelativistic theory equivalent to fully relativistic theory. J Chem Phys, 2005, 123: 241102-1~241102-4.

[79] Filatov M, Dyall K G. On convergence of the normalized elimination of the small component (NESC) method. Theor Chem Acc, 2007, 117: 333~338.

[80] Barysz M, Sadlej A. Two-component methods of relativistic quantum chemistry: from the Douglas-Kroll approximation to the exact two-component formalism. J Mol Struct (Theochem), 2001, 573: 181~200.

[81] 刘文剑. 相对论量子化学新进展. 化学进展, 2007, 19: 833~851.

[82] Heully J L, Lindgren I, Lindroth E, et al. Diagonalisation of the Dirac Hamiltonian as a basis for a relativistic many-body procedure. J Phys B, 1986, 19: 2799~2815.

[83] Liu W J, Kutzelnigg W. Quasirelativistic theory. II. Theory at matrix level. J Chem Phys, 2007, 126: 114107-1~114107-14.

[84] Kutzelnigg W, Liu W J. Quasirelativistic theory. I. Theory in terms of a quasirelativistic operator. Mol Phys, 2006, 104: 2225~2240.

第6章　有效核实势法

一般来说，相对论效应的大小大约和核电荷数的四次方成正比，所以在重元素化学研究中必须考虑相对论效应。由于计算复杂，人们往往只能对原子和小分子作严格的全电子相对论电子结构计算。所以在全电子计算法发展的同时，也发展了许多近似方法。其中能很好地逼近全电子从头算的方法之一就是有效核实势 (effective core potentials，ECPs) 方法。ECPs 方法的主要物理思想是将原子的内层电子连同原子核构成一个核实 (core)，核实的电荷为原子的原子序数 Z 减去内层电子数 n_c。此时核实对外层价电子 i 的作用 (包括相对论效应的贡献) 以某种合适的有效核实势 (ECPs) 来表达。从第 1 章我们已经知道，粒子的速度越接近光速则相对论效应就越明显，由于内层电子的运动速度要大于外层电子，所以一般来说，绝大部分的相对论效应都来自近核区域。既然这部分相对论效应已被包括在 ECPs 中，对外层价电子我们仍可以采用非相对论的 Hamilton 量，进而进行变分和自洽迭代方法计算。这种只对价电子进行严格量子化学处理的方法又称价电子 (valence-only，VO) 方法。与全电子方法相比，价电子方法降低了需要作严格量子化学处理的电子数目，也降低了计算所需基组的规模。随着基组数目的减少，我们也就大大缩短了用于积分计算、HF 或 Kohn-Sham (KS) 自洽场计算、电子相关计算或量子蒙特卡罗计算前的积分变换等过程的机时。

ECPs 主要分为两类，即模型势 (model potentials，MPs) 和赝势 (也称伪势，pseudopotentials，PPs)。在文献中常看到将二者混淆成同一方法的两个不同名称，实际上他们是两种不同的 ECPs 方法。前者可以看成是对核实冻结条件下真实的价电子 HF 算符的模拟，所得到的原子的价电子轨道和真实的全电子轨道的径向节点 (即函数值为零的点) 结构是一样的。后者则是对全电子的价电子轨道进行变换，变换后的轨道称为伪价电子轨道。一般伪价电子轨道的径向节点数目比真实的价电子轨道的径向节点数目少。由于计算所需的基组大小与节点数目成正比，所以 PPs 的配套基组通常要比 MPs 的配套基组更小一些。

大家可能已经注意到，无论是 PPs 方法，还是 MPs 方法，二者都用到了以下两个重要近似:

(1) 内层电子–价层电子分离近似。假定我们可以对核实内电子和核实外电子采用不同的处理方式。

(2) 核实冻结近似。假定核实内的电子分布是不变的，则在特定物理条件下得

到的某原子的 PPs(或 MPs) 可直接应用到该原子与其他原子成键时的情形。

实际上，不仅是 PPs 和 MPs 方法，所有 ECPs 方法都隐含了以上两条近似。需要提醒读者的是，虽然 ECPs 方法是一种用于相对论量子化学计算的近似方法，但并不意味着所有的全电子计算结果都优于 ECPs 的结果，即便是全电子计算选择了更精确的 Hamilton 量。如前所述，与全电子方法相比， ECPs 方法降低了需要严格处理的电子数目，此举不但降低了计算量，还可以对价电子作更为精确的计算，而价电子才是对原子化学性质起决定作用的部分。从第 4 章我们已经知道，在考虑相对论效应后，分子的对称性降低，必须用二倍群来描述，这将导致计算量的大大增加。在 ECPs 框架内，通过选择合适的相对论标量 PPs (scalar-relativistic PPs) 和旋–轨耦合 PPs (SO PPs)，高级别的相关能计算可以仅在自旋平均的相对论级别进行，旋–轨耦合效应则由定义于 LS 或 ΛS 态的有效 Hamilton 算符处理。另外，随着核实极化势 (core-polarization potential，CPP) 的引入，被 ECPs 忽略的核实动态极化效应以及静态极化效应都能得到很好的校正。高精确度的 ECPs 还可以将高级别的相对论效应 (如完整的 Breit 相互作用) 隐含在 ECPs 中，目前分子的全电子计算通常只包含 Breit 效应中的 Gaunt 相互作用。

对计算化学而言，绝大多数原子和分子电子结构从头算使用的都是 Gauss 基组，本章将讨论的从头算相对论 ECPs 也是以 Gauss 函数为表达形式的，介绍的重点是从头算 PPs (*ab initio* PPs) 和从头算 MPs (*ab initio* MPs) 。我们将不涉及其他形式的 ECPs，如固体物理研究中常用的以平面波为展开形式的 ECPs 等。本章内容主要参考了 Klobukowski[1]、Seijo[2]、Pyykkö [3]、Schwerdtfeger[4] 以及作者[5, 6] 本人等关于 ECPs 方法的综述文章，有兴趣的读者可以进一步阅读。

接下来我们将分别介绍 PPs 方法和 MPs 方法的基本概念和重要公式。

6.1　Phillips-Kleinman 方程

早在 1935 年 Hellmann 就将 PPs 方法用于钾的研究[7] (详见 6.5 节)，1959 年 Phillips-Kleinman (PK) 方程[8] 的发表则正式建立了 PPs 的理论基础。该方程由 Phillips 和 Kleinman 导出，用于研究仅含有效单电子的晶体或分子。

假定一个有效的单电子 Hamilton 算符 $\hat{H}_{\mathrm{eff}}$，该 Hamilton 算符仅有一个高能态的本征函数和若干个低能态的本征函数，前者又称价电子本征函数，用 $|\varphi_{\mathrm{v}}\rangle$ 表示，后者又称核实电子本征函数，用 $|\varphi_{\mathrm{c}}\rangle$ 表示。与 $\hat{H}_{\mathrm{eff}}$ 相对应的定态 Schrödinger 方程是

$$\hat{H}_{\mathrm{eff}}|\varphi_a\rangle = \epsilon_a|\varphi_a\rangle \quad (a \in \{\mathrm{v}, \mathrm{c}\}) \tag{6.1}$$

假定所有的本征函数都正交归一

$$\langle\varphi_a|\varphi_{a'}\rangle = \delta_{aa'} \quad (a, a' \in \{\mathrm{v}, \mathrm{c}\}) \tag{6.2}$$

引入新函数 $|\varphi_{\mathrm{p}}\rangle$，$|\varphi_{\mathrm{p}}\rangle$ 定义为 $|\varphi_{\mathrm{c}}\rangle$ 和 $|\varphi_{\mathrm{v}}\rangle$ 的线性组合

$$|\varphi_{\mathrm{p}}\rangle = |\varphi_{\mathrm{v}}\rangle + \sum_{\mathrm{c}} a_{\mathrm{c}}|\varphi_{\mathrm{c}}\rangle \tag{6.3}$$

将上式左乘 $\langle\varphi_{\mathrm{c}}|$，利用本征函数的正交归一性，求得系数 a_{c}

$$a_{\mathrm{c}} = \langle\varphi_{\mathrm{c}}|\varphi_{\mathrm{p}}\rangle \tag{6.4}$$

将 a_{c} 带入 (6.3) 式，我们得到

$$|\varphi_{\mathrm{v}}\rangle = |\varphi_{\mathrm{p}}\rangle - \sum_{\mathrm{c}}\langle\varphi_{\mathrm{c}}|\varphi_{\mathrm{p}}\rangle|\varphi_{\mathrm{c}}\rangle \tag{6.5}$$

再将上式代入 (6.1) 式，由于 $|\varphi_{\mathrm{c}}\rangle$ 也是 $\hat{H}_{\mathrm{eff}}$ 的本征函数，整理得

$$\hat{H}_{\mathrm{eff}}|\varphi_{\mathrm{p}}\rangle - \sum_{\mathrm{c}}\epsilon_{\mathrm{c}}|\varphi_{\mathrm{c}}\rangle\langle\varphi_{\mathrm{c}}|\varphi_{\mathrm{p}}\rangle = \epsilon_{\mathrm{v}}\left[|\varphi_{\mathrm{p}}\rangle - \sum_{\mathrm{c}}|\varphi_{\mathrm{c}}\rangle\langle\varphi_{\mathrm{c}}|\varphi_{\mathrm{p}}\rangle\right] \tag{6.6}$$

定义“Phillips-Kleinman (PK) 势”$\hat{V}^{\mathrm{PK}}$

$$\hat{V}^{\mathrm{PK}} = \sum_{\mathrm{c}}(\epsilon_{\mathrm{v}} - \epsilon_{\mathrm{c}})|\varphi_{\mathrm{c}}\rangle\langle\varphi_{\mathrm{c}}| \tag{6.7}$$

将 $\hat{V}^{\mathrm{PK}}$ 代入 (6.6) 式，不难得到

$$\left(\hat{H}_{\mathrm{eff}} + \hat{V}^{\mathrm{PK}}\right)|\varphi_{\mathrm{p}}\rangle = \epsilon_{\mathrm{v}}|\varphi_{\mathrm{p}}\rangle \tag{6.8}$$

上式就是 PK 方程。由 PK 方程我们发现，与原价电子本征函数 $|\varphi_{\mathrm{v}}\rangle$ 对应的能量本征值 ϵ_{v} 也是与新定义函数 $|\varphi_{\mathrm{p}}\rangle$ 对应的能量本征值。由于 $|\varphi_{\mathrm{p}}\rangle$ 并不是真实的 Hamilton 算符 $\hat{H}_{\mathrm{eff}}$ 的本征函数，所以 $|\varphi_{\mathrm{p}}\rangle$ 又称为伪价电子本征函数。

由 (6.3) 式，如果选择恰当的 $|\varphi_{\mathrm{v}}\rangle$ 和 $|\varphi_{\mathrm{c}}\rangle$ 的线性组合，可使伪价电子本征函数 $|\varphi_{\mathrm{p}}\rangle$ 的节点数目少于 $|\varphi_{\mathrm{v}}\rangle$，则变分求解 PK 方程 (6.8) 式的计算量将小于求解 (6.1) 式的计算量。另外，求解 PK 方程需要知道 PK 势 $\hat{V}^{\mathrm{PK}}$，由 (6.7) 式知构造 PK 势需要知道 $|\varphi_{\mathrm{c}}\rangle$ 和 $\epsilon_a(a=\mathrm{v},\mathrm{c})$。也就是说，与求解真实的 Hamilton 方程 (6.1) 式相比求解 PK 方程并不容易。并且，从 (6.3) 式我们可以构造多个与 $|\varphi_{\mathrm{v}}\rangle$ 有相同能量本征值的伪价电子本征函数 $|\varphi_{\mathrm{p}}\rangle$，在对研究化学性质很重要的价电子空间区域，这些 $|\varphi_{\mathrm{p}}\rangle$ 很可能与真实的价电子本征函数 $|\varphi_{\mathrm{v}}\rangle$ 差别很大，常导致较大的计算误差。关于 PK 理论研究的最新进展读者可以阅读 Schwartz 等的文章[9, 10]。

本节讨论的 PK 方程仅适用于有效单电子体系，接下来我们讨论适用于多电子体系的 PK 方程。

6.2 广义的 Phillips-Kleinman 方程

1968 年，Weeks 和 Rice[11] 将 PK 方程推广到多电子体系，导出广义的 Phillips-Kleinman 方程。

假定一个代表 n_v 个价电子的哈密顿算符 $\hat{H}_\mathrm{v}$，一套正交归一的核实内电子的本征函数 $|\varphi_\mathrm{c}\rangle$。我们希望求下列积分

$$\langle \Phi_\mathrm{v}|\hat{H}_\mathrm{v}|\Phi_\mathrm{v}\rangle \tag{6.9}$$

的极小值。$|\Phi_\mathrm{v}\rangle$ 是归一化的，且和核实内电子的本征函数 $|\varphi_\mathrm{c}\rangle$ 正交，即

$$\langle \Phi_\mathrm{v}|\Phi_\mathrm{v}\rangle = 1 \tag{6.10}$$

$$\langle \Phi_\mathrm{v}|\varphi_\mathrm{c}\rangle = 0 \tag{6.11}$$

定义投影算符 $\hat{P}$

$$\hat{P} = \sum_\mathrm{c} |\varphi_\mathrm{c}\rangle\langle\varphi_\mathrm{c}| \tag{6.12}$$

$\hat{P}$ 是幂等厄米算符 (见附录 A)，容易证明 $\hat{P}$ 满足

$$\begin{aligned} \hat{P}^2 &= \hat{P} \\ \hat{P}|\varphi_\mathrm{c}\rangle &= |\varphi_\mathrm{c}\rangle \\ (1-\hat{P})|\varphi_\mathrm{c}\rangle &= 0 \end{aligned} \tag{6.13}$$

引入伪价电子本征函数 $|\Phi_\mathrm{p}\rangle$，将投影算符 $\hat{P}$ 作用到 $|\Phi_\mathrm{p}\rangle$ 即得到 $|\varphi_\mathrm{c}\rangle$，我们将 $|\Phi_\mathrm{v}\rangle$ 重写为

$$|\Phi_\mathrm{v}\rangle = (1-\hat{P})|\Phi_\mathrm{p}\rangle \tag{6.14}$$

将上式中的 $|\Phi_\mathrm{v}\rangle$ 代回 (6.9) 式，则我们将求 (6.9) 式的极小值问题转化为求下列积分的极小值

$$\langle (1-\hat{P})\Phi_\mathrm{p}|\hat{H}_\mathrm{v}|(1-\hat{P})\Phi_\mathrm{p}\rangle \tag{6.15}$$

归一化条件为

$$\langle (1-\hat{P})\Phi_\mathrm{p}|(1-\hat{P})\Phi_\mathrm{p}\rangle = 1 \tag{6.16}$$

采用 Lagrange 乘子法 (见附录 F)，求 (6.15) 式的极小值问题可转化为求以下函数 F 的极小值

$$F = \langle (1-\hat{P})\Phi_\mathrm{p}|\hat{H}_\mathrm{v}|(1-\hat{P})\Phi_\mathrm{p}\rangle - E_\mathrm{v}\left[\langle (1-\hat{P})\Phi_\mathrm{p}|(1-\hat{P})\Phi_\mathrm{p}\rangle - 1\right] \tag{6.17}$$

F 取极小值条件是 F 的一阶导数为零，将上式对 $\langle\Phi_\mathrm{p}|$ 求导得

$$\frac{\partial F}{\partial \Phi_\mathrm{p}^*} = \left(1-\hat{P}\right)\hat{H}_\mathrm{v}\left(1-\hat{P}\right)|\Phi_\mathrm{p}\rangle - E_\mathrm{v}\left(1-\hat{P}\right)^2|\Phi_\mathrm{p}\rangle = 0 \tag{6.18}$$

利用 (6.13) 式，将上式整理成与 (6.8) 式类似的形式

$$\left(\hat{H}_{\rm v}+\hat{V}^{\rm GPK}\right)|\Phi_{\rm p}\rangle=E_{\rm v}|\Phi_{\rm p}\rangle \tag{6.19}$$

上式即为广义的 (generalized) Phillips-Kleinman (GPK) 方程。方程中的 $\hat{V}^{\rm GPK}$ 定义为广义的 Phillips-Kleinman 势

$$\hat{V}^{\rm GPK}=-\hat{H}_{\rm v}\hat{P}-\hat{P}\hat{H}_{\rm v}+\hat{P}\hat{H}_{\rm v}\hat{P}+E_{\rm v}\hat{P} \tag{6.20}$$

由于 $\hat{V}^{\rm GPK}$ 并非真正的势，所以称为赝势 (PPs)。由上可以看出，通过引入伪价电子本征函数 $|\Phi_{\rm p}\rangle$，我们将给定体系的 Hamilton 方程求解过程转化为求解 GPK 方程，即求解 (6.19) 式和 (6.20) 式。显然，由于 GPK 方程只是原方程的另一种更为复杂的表达形式，所以 GPK 方程的求解过程并不比求解 Hamilton 算符简单。但是，正如 Weeks 和 Rice 指出 —— “PPs 方法带来的便利并不在于 PPs 的精确求解过程，而在于该方法所揭示的物理现象和引入的物理模型”。PK 和 GPK 方程告诉我们，可以将一个人造的 PPs 引入到真实的价电子 Hamilton 算符中从而得到价电子模型 Hamilton 算符，此价电子模型 Hamilton 算符和真实的价电子 Hamilton 算符的能量本征值相同，并且价电子模型 Hamilton 算符的伪价电子本征函数 $|\Phi_{\rm p}\rangle$ 并不需要和核实内电子的本征函数 $|\varphi_{\rm c}\rangle$ 正交。

PK 和 GPK 方程为现代 PPs 的产生和发展提供了坚实的理论基础，接下来我们讨论符合 GPK 方程的原子的价电子模型 Hamilton 算符。

6.3 原子的价电子模型 Hamilton 算符

设原子核的电荷为 Z，核实内的电子数为 $n_{\rm c}$，核实外价电子数目为 $n_{\rm v}$。在固定核实近似下，原子的价电子 Hamilton 算符 $\hat{H}_{\rm v}$ 是

$$\hat{H}_{\rm v}=-\frac{1}{2}\sum_{i}^{n_{\rm v}}\nabla_i^2+\sum_{i<j}^{n_{\rm v}}\hat{g}(i,j)-\sum_{i}^{n_{\rm v}}\frac{Q}{r_i}+\sum_{i}^{n_{\rm v}}\hat{V}'_{\rm cv}(i) \tag{6.21}$$

式中，$\hat{g}(i,j)$ 代表电子–电子相互作用，通常写成 Coulomb 相互作用形式 —— $\hat{g}(i,j)=1/r_{ij}$；Q 是核实所带电荷，$Q=Z-n_{\rm c}$；$\hat{V}'_{\rm cv}(i)$ 代表电子 i 和核实之间的相互作用，其主要贡献为 Coulomb 相互作用 $-Q/r_i$。根据 (6.19) 式，把 GPK 势 $\hat{V}^{\rm GPK}$ 引入 $H_{\rm v}$，得

$$\hat{H}_{\rm v}^{\rm GPK}=-\frac{1}{2}\sum_{i}^{n_{\rm v}}\nabla_i^2+\sum_{i<j}^{n_{\rm v}}\hat{g}(i,j)-\sum_{i}^{n_{\rm v}}\frac{Q}{r_i}+\sum_{i}^{n_{\rm v}}\hat{V}'_{\rm cv}(i)+\hat{V}^{\rm GPK} \tag{6.22}$$

使用独立粒子模型，将 $\hat{V}^{\rm GPK}$ 写成所有单个电子的 $\hat{V}^{\rm GPK}(i)$ 之和，则上式中的后两项可合并为

$$\Delta\hat{V}_{\rm cv}(i)=\hat{V}'_{\rm cv}(i)+\hat{V}^{\rm GPK}(i) \tag{6.23}$$

对价电子 i，定义一个有效的单电子势

$$\hat{V}_{\rm cv}(i) = -\frac{Q}{r_i} + \Delta\hat{V}_{\rm cv}(i) \tag{6.24}$$

利用上式，将 (6.22) 式重写为

$$\hat{H}_{\rm v}^{\rm PP} = -\frac{1}{2}\sum_{i}^{n_{\rm v}} \nabla_i^2 + \sum_{i<j}^{n_{\rm v}} \hat{g}(i,j) + \sum_{i}^{n_{\rm v}} \hat{V}_{\rm cv}(i) \tag{6.25}$$

$\hat{H}_{\rm v}^{\rm PP}$ 称为原子的价电子模型 Hamilton 算符，式中的 $\hat{V}_{\rm cv}(i)$ 定义为赝势 (PPs)，代表了价电子 i 和核实的相互作用。在 6.5 节和 6.6 节我们将讨论 PPs 的具体形式。

6.4　分子的价电子模型 Hamilton 算符

在核实冻结近似下，由 $N_{\rm c}$ 个核实组成的共有 $n_{\rm v}$ 个价电子的分子体系的模型 Hamilton 算符可以写成

$$\hat{H}_{\rm v} = -\frac{1}{2}\sum_{i}^{n_{\rm v}} \nabla_i^2 + \sum_{i<j}^{n_{\rm v}} \hat{g}(i,j) - \sum_{i}^{n_{\rm v}} \frac{Q}{r_i} + \sum_{i}^{n_{\rm v}} \hat{V}'_{\rm cv}(i) + V_{\rm cc} \tag{6.26}$$

与原子的 Hamilton 算符相比，分子的 Hamilton 算符增加了核实–核实相互作用项 $V_{\rm cc}$，该项可写成以下两项的和

$$V_{\rm cc} = \sum_{\lambda<\mu}^{N_{\rm c}} \left[\frac{Q_\lambda Q_\mu}{r_{\lambda\mu}} + \Delta V_{\rm cc}^{\lambda\mu}(r_{\lambda\mu})\right] \tag{6.27}$$

$V_{\rm cc}$ 的主要贡献来自点电荷模型下的 Coulomb 排斥项 (第一项)，第二项为点电荷模型校正项，我们将在 6.8 节详细讨论第二项的具体函数形式。

采用与 6.3 节类似的步骤 (见 (6.23) 式 ~(6.25) 式)，不难得到分子的价电子模型 Hamilton 算符

$$\hat{H}_{\rm v} = -\frac{1}{2}\sum_{i}^{n_{\rm v}} \nabla_i^2 + \sum_{i<j}^{n_{\rm v}} \hat{g}(i,j) + \sum_{i}^{n_{\rm v}} \hat{V}_{\rm cv}(i) + V_{\rm cc} \tag{6.28}$$

其中

$$\hat{V}_{\rm cv}(i) = -\sum_{\lambda}^{N_{\rm c}} \frac{Q_\lambda}{r_{\lambda i}} + \Delta\hat{V}_{\rm cv}(i) \tag{6.29}$$

由核实冻结近似，可以认为各个核实是互不重叠的，则分子的赝势 $\hat{V}_{\rm cv}(i)$ 可以写成对所有单个核实的赝势求和

$$\hat{V}_{\rm cv}(i) = \sum_{\lambda}^{N_{\rm c}} \hat{V}_{\rm cv}^{\lambda}(i) = \sum_{\lambda}^{N_{\rm c}} \left[-\frac{Q_\lambda}{r_{\lambda i}} + \Delta\hat{V}_{\rm cv}^{\lambda}(i)\right] \tag{6.30}$$

下面我们接着讨论 $\Delta\hat{V}_{\rm cv}$ 和 $\Delta V_{\rm cc}$ 的函数形式。

6.5 非相对论赝势的解析式

历史上第一个赝势函数是 Hellmann 于 1935 年提出的[7]，他将 19 号元素钾处理为含 18 个内层电子 (1s~3p) 的核实和一个价电子 (4s)，价电子与核实的相互作用以一个简单的径向 (又称局域，local) 函数来表达

$$\hat{V}_{\rm cv}(\boldsymbol{r}) = V_{\rm cv}(r) = -\frac{1}{r} + \frac{2.74}{r}{\rm e}^{-1.16r} \tag{6.31}$$

很快，人们就发现仅用径向函数表达的赝势 (又称局域赝势) 常带来很大的误差[12]。1965 年，在固体研究中， Abarenkov 和 Heine 首先将角分布函数引入到赝势的解析式中[13]。三年后，Schwarz[14]，Kahn 和 Goddard[15] 等将此思想也运用到量子化学计算中。1972 年，在前人工作的基础上 Kahn 和 Goddard 提出了统一的非相对论赝势算符[16]

$$\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i) = \sum_{l,m} V_l(r_i)|lm\rangle\langle lm| = \sum_{l=0}^{\infty} V_l(r_i)\hat{P}_l(i) \tag{6.32}$$

式中，$\hat{P}_l(i)$ 是用球谐函数 $|lm\rangle$ 展开的角动量投影算符

$$\hat{P}_l(i) = \sum_{m=-l}^{l} |lm\rangle\langle lm| \tag{6.33}$$

Kahn 和 Goddard 发现，随着 l 的增大，$V_l(r_i)$ 的变化逐渐减少，当 l 大于某一值 L 时，通常 $(L-1)$ 是核实内电子轨道的最高角量子数，(6.32) 式中的 $V_l(r_i)$ 变化很小，可以近似认为

$$V_l(r_i) \cong V_L(r_i) \qquad (l \geqslant L) \tag{6.34}$$

于是他们假定 $V_l(r_i) = V_L(r_i)$ $(l \geqslant L)$，利用投影算符的封闭性 (见附录 A)，将 (6.32) 式改写为

$$\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i) \cong V_L(r_i) + \sum_{l=0}^{L-1}\Big[V_l(r_i) - V_L(r_i)\Big]\hat{P}_l(i) \tag{6.35}$$

如果选择 (6.32) 式中的 $V_l(r_i)$ 仅是 r 的函数，与 l 无关，则可将 $V_l(r_i) = V(r_i)$ 提到求和号外，即得局域赝势算符

$$\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i) = V(r_i)\sum_{l=0}^{\infty}\hat{P}_l(i) = V(r_i) \tag{6.36}$$

(6.35) 式定义的赝势，仅对大于 $(L-1)$ 的 l 值时取相同径向函数，当角量子数 $0 \leqslant l \leqslant (L-1)$ 时，每一个角量子数 l 对应一个不同的径向函数 $(V_l(r_i) - V_L(r_i))$，所以 (6.35) 式定义的赝势又称为半局域赝势。

为方便计算包含赝势算符的矩阵元，Kahn, Baybutt 和 Truhlar[17] 将 $V_l(r_i)$ 和 $V_L(r_i)$ 写成类似 Gauss 函数的线性组合形式

$$V_m(r_i) = \sum_k A_{km} r_i^{n_{km}} \mathrm{e}^{-a_{km} r_i^2} \qquad (m = l, L) \tag{6.37}$$

式中，r_i 代表电子和核实的距离。为获得展开式中的系数 A_{km} 和指数 a_{km}，Kahn，Baybutt 和 Truhlar[17] 提出了如下步骤来构造非相对论从头算赝势。

第一步, 对每一个角量子数 $l(l = 0, 1, \cdots, L,\ L-1$ 是核实内电子轨道的最大角量子数)，将对应的伪价电子轨道 ($|\varphi_\mathrm{p}\rangle$) 按 (6.3) 式写成真实的价电子轨道和核实内电子轨道的线性组合。线性组合系数优化必须满足以下约束条件：

(1) 对给定的角量子数 l，对应的能量最低的伪价电子轨道 $|\varphi_\mathrm{p}\rangle$ 在 $r > 0$ 的区域是无节点的。

(2) 伪价电子轨道 $|\varphi_\mathrm{p}\rangle$ 要尽量逼近真实的价电子轨道 $|\varphi_\mathrm{v}\rangle$。

(3) 伪价电子轨道 $|\varphi_\mathrm{p}\rangle$ 的空间曲线要尽量光滑。

按以上条件优化得到的 $|\varphi_\mathrm{p}\rangle$，其节点结构比真实的价电子轨道相对简单，精确表达 $|\varphi_\mathrm{p}\rangle$ 所需的赝势基组也就小于描述真实价电子轨道所需的全电子基组。

第二步, 把第一步得到的 $|\varphi_\mathrm{p}\rangle$ 和对应的价电子轨道能量 ϵ_v 代入径向 Fock 方程，再用数值法反演求解赝势 ($V_l(r_i)$) 。最后得到函数值表，即以表格形式给出当 r 取某些值时相应的 $V_l(r_i)$ 的数值。

第三步，用最小二乘法，将第二步得到的数值赝势 ($V_l(r_i)$) 根据 (6.37) 式进行拟合，最后得到展开系数和指数。关于 r 的指数 (n_{km})，Kahn 等发现，当限制其取值范围为 -2, -1, 0 时，所得赝势函数解析形式就可以正确逼近赝势在原点 ($r = 0$) 的表现[17]。

在与化学反应活性相关的价电子空间区域，按照以上方法第一步得到的伪价电子轨道和真实的价电子轨道有时差别很大，该误差被带入第二步，结果影响赝势的精度。如果将此赝势用于量子化学计算甚至可能导致完全错误的结论。尽管如此，Kahn 等的研究为从头算赝势的发展起到了很重要的作用，如他们提出的赝势的 Gauss 函数展开形式、指数 n_{km} 的取值范围等，至今仍被现代赝势广泛使用。

为解决在价电子空间区域伪价电子轨道 ($|\varphi_\mathrm{p}\rangle$) 和真实的价电子轨道 ($|\varphi_\mathrm{v}\rangle$) 差别太大的问题，Durand 和 Barthelat[18] 以及 Christiansen, Lee, Pitzer[19] 等先后提出了形状一致赝势 (shape-consistent PPs) 的概念 —— 要求 $|\varphi_\mathrm{p}\rangle$ 和 $|\varphi_\mathrm{v}\rangle$ 在价电子空间区域一致，我们将在 6.9.2 节详细讨论形状一致的从头算赝势。

6.6　相对论赝势的解析式

读者可能已经注意到 Kahn 等的赝势形式 (见 (6.32) 式) 是基于角动量量子数

(l) 展开的，并没有考虑自旋。他们选择的全电子参考轨道和能量也是非相对论的计算结果，也就是说没有考虑任何的相对论效应。实际上，我们有两个途径来考虑相对论效应。一个途径是保留 Kahn 等的赝势形式不变，而将参考的全电子价电子轨道和能量换成全电子相对论自旋平均 (或忽略自旋) 计算 (即一分量计算) 的结果，所得赝势称为相对论标量赝势 (scalar-relativistic PPs) 或一分量赝势。另一个途径是将总量子数 (j) 引入赝势函数展开，用包含了自旋的全电子二分量计算来得到参考态数据，所得相对论赝势又称为二分量赝势。

1978 年，Kahn, Hay 和 Cowan 发表了第一个相对论标量赝势[20]，他们选择了全电子 Cowan-Griffin (CG)[21] Hamilton 算符 (见第 5 章 5.2.4 节) 计算参考态数据。几乎同时，Hafner 和 Schwarz 发表了第一个半经验二分量赝势[22]，Lee, Ermler 和 Pitzer 发表了第一个从头算二分量赝势[23]。接下来我们就来介绍从头算二分量赝势的解析形式。

在非相对论的情形，具有相同主量子数 (n) 和角量子数 (l) 的 HF 轨道的能量是简并的，相应的半局域赝势可写成关于角量子数 (l) 的函数展开 (见 (6.32) 式)。在考虑自旋的条件下，由于简并度的降低我们将半局域赝势算符写成关于 l 和 j 的函数展开

$$\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i) = \sum_{l,j,m} V_{lj}(r_i)|ljm\rangle\langle ljm| = \sum_{l=0}^{\infty}\sum_{j=|l-\frac{1}{2}|}^{l+\frac{1}{2}} V_{lj}(r_i)\hat{P}_{lj}(i) \tag{6.38}$$

投影算符 $\hat{P}_{lj}$ 由自旋量球谐函数 $|ljm\rangle$ 构成

$$\hat{P}_{lj}(i) = \sum_{m=-j}^{j} |ljm\rangle\langle ljm| \tag{6.39}$$

采用与 Kahn 和 Goddard 方法相同的步骤 (见 6.5 节)，容易得到 $\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i)$

$$\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i) \cong V_{LJ} + \sum_{l=0}^{L-1}\sum_{j=|l-\frac{1}{2}|}^{l+\frac{1}{2}}\Big[V_{lj}(r_i) - V_{LJ}(r_i)\Big]\hat{P}_{lj} \tag{6.40}$$

利用以下条件

$$\hat{P}_l(i) = \sum_j \hat{P}_{lj}(i) = \hat{P}_{l,|l-\frac{1}{2}|}(i) + \hat{P}_{l,l+\frac{1}{2}}(i) \tag{6.41}$$

$$V_l(r_i) = \frac{1}{2l+1}\Big[lV_{l,|l-\frac{1}{2}|}(r_i) + (l+1)V_{l,l+\frac{1}{2}}(r_i)\Big] \tag{6.42}$$

我们可以将相对论赝势 (6.40) 式写成两项的和，一项为自旋平均项 (av)，另一项为旋–轨耦合项 (so)[24]

$$\Delta\hat{V}_{\rm cv}(\boldsymbol{r}_i) = \Delta\hat{V}_{\rm cv,av}(\boldsymbol{r}_i) + \Delta\hat{V}_{\rm cv,so}(\boldsymbol{r}_i) \tag{6.43}$$

其中

$$\Delta\hat{V}_{\text{cv,so}}(\boldsymbol{r}_i)=\sum_{l=1}^{L-1}\frac{\Delta V_l(r_i)}{2l+1}\left[l\hat{P}_{l,l+\frac{1}{2}}(i)-(l+1)\hat{P}_{l,|l-\frac{1}{2}|}(i)\right] \tag{6.44}$$

$$\Delta\hat{V}_l(r_i)=V_{l,l+\frac{1}{2}}-V_{l,|l-\frac{1}{2}|} \tag{6.45}$$

$\Delta\hat{V}_{\text{cv,av}}(\boldsymbol{r}_i)$ 和 (6.35) 式定义的 $\Delta\hat{V}_{\text{cv}}(\boldsymbol{r}_i)$ 一致[24]。由于 $\Delta\hat{V}_{\text{cv,av}}(\boldsymbol{r}_i)$ 中不包含旋–轨耦合，所以称为相对论标量赝势 (scalar-relativistic PPs)。相对论标量赝势也称一分量赝势，与此相对应的包含旋–轨耦合的相对论赝势称为二分量赝势。由上可以看出，相对论 (即二分量) 赝势可以自旋平均化后用于相对论标量 (即一分量) 计算，$\Delta\hat{V}_{\text{cv,so}}(\boldsymbol{r}_i)$ 称为旋–轨耦合赝势，下面我们将其写为更常用的形式。

由于

$$\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}=\frac{1}{2}\left(\hat{\boldsymbol{J}}^2-\hat{\boldsymbol{L}}^2-\hat{\boldsymbol{S}}^2\right) \tag{6.46}$$

可得下列积分

$$\begin{aligned}
&\langle l,l+\frac{1}{2},m_j|\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}|l,l+\frac{1}{2},m_j'\rangle=\frac{l}{2}\delta_{m_jm_j'}\\
&\langle l,l+\frac{1}{2},m_j|\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}|l,l-\frac{1}{2},m_j'\rangle=0\qquad(l>1)\\
&\langle l,l-\frac{1}{2},m_j|\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}|l,l-\frac{1}{2},m_j'\rangle=-\frac{(l+1)}{2}\delta_{m_jm_j'}\qquad(l>1)
\end{aligned} \tag{6.47}$$

(6.44) 式中方括号内的两项可改写为

$$\begin{aligned}
l\hat{P}_{l,l+\frac{1}{2}}-(l+1)\hat{P}_{l,l-\frac{1}{2}}&=2\sum_{jj'm_jm_j'}|l,j,m_j\rangle\langle l,j,m_j|\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}|l,j',m_j'\rangle\langle l,j',m_j'|\\
&=2\sum_{jj'}\hat{P}_{lj}\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}\hat{P}_{lj'}\\
&=2\hat{P}_l\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}}\hat{P}_l
\end{aligned} \tag{6.48}$$

代入 (6.44) 式，得

$$\Delta\hat{V}_{\text{cv,so}}(\boldsymbol{r}_i)=\sum_{l=1}^{L-1}\frac{2\Delta V_l(r_i)}{2l+1}\hat{P}_l(i)\hat{\boldsymbol{L}}_i\cdot\hat{\boldsymbol{S}}_i\hat{P}_l(i) \tag{6.49}$$

上式由 Pitzer 和 Winter 导出[25]，常用于包含旋–轨耦合作用的组态相互作用 (spin-orbit configuration interaction，SOCI) 计算。式中 $\boldsymbol{L}_i=\hat{\boldsymbol{r}}_i\times\hat{\boldsymbol{p}}_i$ 和 $\hat{\boldsymbol{S}}_i$ 分别代表轨道角动量和自旋角动量算符。

在作 SOCI 计算时，由于分子对称性的降低，必须以相对论二倍群取代非相对论的点群，此时的 CI 又称为二倍群 CI (double group CI，DGCI)。据测算，对给定

空间组态，用于二分量计算的 Slater 行列式的数目约为一分量计算的六倍，所以高级别的直接 DGCI 计算 (又称一步法) 是很困难的。为减少计算量，人们提出了先考虑相关能再考虑旋–轨耦合的二步法，即先在自旋平均水平作高精度的 CI 计算 (一分量计算)，然后再作旋–轨耦合处理 (可以是微扰或 CI 方法)。常见的一步法有 Christiansen 等的 CIDBG (全称：CI calculation using double-group-adapted configuration functions)[26, 27], DiLabio 等的 SICCI(全称：selected intermediate coupling CI)[27]，Esser[28]、Sjovoll 等[29] 的 LUCIA (全称：Lund CI approach) 等。常见的二步法有 Vallet 等的 EPCISO (全称：effective and polarized CI-SO)[30], Hess 等的平均场旋–轨耦合 (mean-field-SO) 法[31] 等。

需要提醒读者的是，与自旋平均的相对论赝势配套的旋–轨耦合赝势通常仅对价电子轨道优化，没有考虑半壳层轨道，适用于将旋–轨耦合作微扰处理，如果将旋–轨耦合赝势用于二分量的变分计算常导致较大的误差。

6.7 核实极化势

以上讨论的原子 (分子) 的价电子模型 Hamilton 算符使用的是独立粒子模型。从价电子轨道出发我们建立了 HF 形式的行列式波函数 —— 也就是说我们并没有考虑电子相关效应。在优化赝势参数时，多数情况下目标函数来自参考态的全电子 HF 计算结果。当核实被局域 (或半局域) 赝势函数取代后，我们可以对核实外电子作严格的电子相关处理，但显然我们只能忽略核实外电子和核实内电子、以及核实内电子之间的电子相关，此相关效应对核实内电子分布的影响称为动态核实极化。全电子的组态相互作用 (CI) 计算表明，内层轨道的单激发态和价层轨道电子跃迁至高能量空轨道的单激发态之间的相互作用，是价层和内层电子相关能的主要来源。

与其他有效核实势方法一样，赝势方法采用了核实冻结近似，这一近似是 Born-Oppenheimer 近似 (又称固定核近似) 的延伸，即假定核实内的原子核和电子分布是固定的，这里我们没有考虑核实外电子对核实的极化作用 (又称静态核实极化)。

经验表明，在某些情况下核实冻结近似会给计算结果带来较大的误差。例如，当对第 1, 2, 11, 12 族元素使用大核实赝势，即核实电子数很多而价电子数很少时，或者是核实本身易于极化时。此时有必要在价电子模型 Hamilton 算符 (见 (6.28) 式) 中引入核实极化势 (core-polarization potential，CPP) $\hat{V}_{\text{cpp}}$

$$\hat{H}_{\text{v}} = -\frac{1}{2}\sum_{i}^{n_{\text{v}}} \nabla_i^2 + \sum_{i<j}^{n_{\text{v}}} \hat{g}(i,j) + \sum_{i}^{n_{\text{v}}} \hat{V}_{\text{cv}}(i) + V_{\text{cc}} + \hat{V}_{\text{cpp}} \tag{6.50}$$

利用电动力学中电荷体系在外电场中能量的多级展开法可求得核实极化势的解析

形式。

如图 6.1 所示，以原子核为原点 O，设核实 (用 V' 表示) 内电荷分布密度为 $\rho(\boldsymbol{r}')$，则核实外一点 A (用 $\boldsymbol{r}$ 表示) 的标势可写为

$$\varphi(\boldsymbol{r})=\int_{V'}\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}V' \tag{6.51}$$

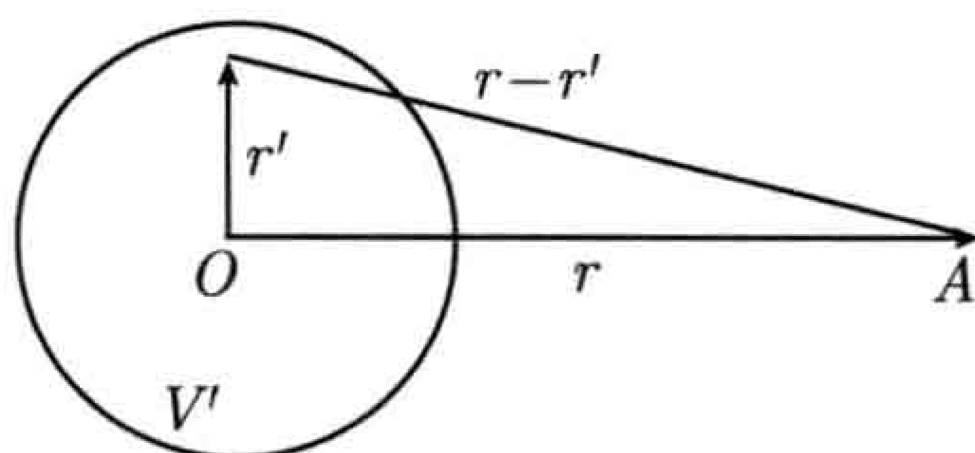


图 6.1　核实 V' 和核实外一点 A

当 $|\boldsymbol{r}|\gg|\boldsymbol{r}'|$ 时，将 $\dfrac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}$ 作级数展开

$$\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}=\frac{1}{|\boldsymbol{r}|}-\boldsymbol{r}'\cdot\nabla\frac{1}{|\boldsymbol{r}|}+\frac{1}{2}\sum_{ij}r_i'r_j'\frac{\partial^2}{\partial r_i\partial r_j}\frac{1}{|\boldsymbol{r}|}+\cdots \tag{6.52}$$

式中，$i,j=x,y,z$。将上式代入 (6.51) 式，整理得

$$\varphi(\boldsymbol{r})=\frac{Q}{r}-\boldsymbol{P}\cdot\nabla\frac{1}{r}+\frac{1}{6}\sum_{ij}D_{ij}\frac{\partial^2}{\partial r_i\partial r_j}\frac{1}{r}+\cdots \tag{6.53}$$

式中，Q, $\boldsymbol{P}$ 和 $D_{ij}(i,j=x,y,z)$ 分别代表核实所带电荷、电偶极矩和电四极矩

$$\begin{aligned}Q&=\int_{V'}\rho(\boldsymbol{r}')\mathrm{d}V'\\ \boldsymbol{P}&=\int_{V'}\rho(\boldsymbol{r}')\boldsymbol{r}'\mathrm{d}V'\\ D_{ij}&=\int_{V'}3r_i'r_j'\rho(\boldsymbol{r}')\mathrm{d}V'\qquad(i,j=x,y,z)\end{aligned} \tag{6.54}$$

设在点 A 处有一电荷 q，则核实 Q 与电荷 q 的相互作用能为

$$W=q\varphi(\boldsymbol{r})\approx\frac{qQ}{r}-\boldsymbol{P}\cdot q\nabla\frac{1}{r} \tag{6.55}$$

根据 Fuentealba 等的研究，只需要考虑核实内电荷在外场中的偶极极化就能起到很好的校正作用[32]，所以上式我们略去了电四极矩及其他高阶展开项。

注意到 (6.55) 式中 $q\nabla\dfrac{1}{r}$ 即为电荷 q 在核实的电场强度，用 $\boldsymbol{E}_{\text{ext}}$ 表示。由于静电作用，外场 $\boldsymbol{E}_{\text{ext}}$ 的存在会影响核实电荷的分布，从而影响电偶极矩，我们可

以把核实的电偶极矩写为永久电偶极矩 $\boldsymbol{P}_0$ 与由 $\boldsymbol{E}_{\text{ext}}$ 产生的诱导偶极矩 $\boldsymbol{\mu}$ 两项之和

$$\begin{aligned}\boldsymbol{P} &= \boldsymbol{P}_0 + \boldsymbol{\mu} \\ \boldsymbol{\mu} &= \alpha_{\text{c}} \boldsymbol{E}_{\text{ext}}\end{aligned} \tag{6.56}$$

式中，α_{c} 为核实的极化率。于是 (6.55) 式可重写为

$$W = q\varphi(\boldsymbol{r}) \approx \frac{qQ}{r} - \boldsymbol{P}_0 \cdot \boldsymbol{E}_{\text{ext}} - \alpha_{\text{c}} \boldsymbol{E}_{\text{ext}}^2 \tag{6.57}$$

上式第三项即为我们寻找的由于核实极化引起的能量变化。

对含多个原子的分子，在外电场 $\hat{\boldsymbol{f}}_\lambda(\boldsymbol{r})$ 中，核实 λ 的诱导偶极矩 $\hat{\boldsymbol{\mu}}_\lambda$ 可写为

$$\hat{\boldsymbol{\mu}}_\lambda = \alpha_\lambda \hat{\boldsymbol{f}}_\lambda \tag{6.58}$$

$\hat{\boldsymbol{f}}_\lambda(\boldsymbol{r})$ 等于所有其他核实、核、以及价电子在核实 λ 处产生的电场叠加

$$\hat{\boldsymbol{f}}_\lambda = -\sum_i \frac{\boldsymbol{r}_{i\lambda}}{r_{i\lambda}^3} + \sum_{\mu \neq \lambda} Q_\mu \frac{\boldsymbol{r}_{\mu\lambda}}{r_{\mu\lambda}^3} \tag{6.59}$$

式中，Q_μ 代表核实 (或核) μ 所带的电荷。分子核实极化势定义为诱导偶极矩和外电场的相互作用能

$$\hat{V}_{\text{cpp}} = -\frac{1}{2}\sum_\lambda \hat{\boldsymbol{\mu}}_\lambda \cdot \hat{\boldsymbol{f}}_\lambda = -\frac{1}{2}\sum_\lambda \alpha_\lambda \hat{\boldsymbol{f}}_\lambda^2 \tag{6.60}$$

为避免 $r \to 0$ 时上式出现发散的情形， Müller, Flesch 和 Meyer 引入了截断函数 F，他们将核实 λ 处的外场 $\hat{\boldsymbol{f}}_\lambda$ 重写为[33]

$$\hat{\boldsymbol{f}}_\lambda = -\sum_i \frac{\boldsymbol{r}_{i\lambda}}{r_{i\lambda}^3} F(r_{i\lambda}, \delta_{\text{e}}^\lambda) + \sum_{\mu \neq \lambda} Q_\mu \frac{\boldsymbol{r}_{\mu\lambda}}{r_{\mu\lambda}^3} F(r_{\mu\lambda}, \delta_{\text{c}}^\lambda) \tag{6.61}$$

截断函数 F 的常见形式是

$$F(r_{i\lambda}, \delta_{\text{e}}^\lambda) = \left[1 - \exp(-\delta_{\text{e}}^\lambda r_{i\lambda}^2)\right]^{n_{\text{e}}} \tag{6.62}$$

$$F(r_{\mu\lambda}, \delta_{\text{c}}^\lambda) = \left[1 - \exp(-\delta_{\text{c}}^\lambda r_{\mu\lambda}^2)\right]^{n_{\text{c}}} \tag{6.63}$$

式中，指数 n_{e} 和 n_{c} 一般选为 1 或 2；截断参数 $\delta_{\text{e}}^\lambda$ 和 $\delta_{\text{c}}^\lambda$ 可通过优化求得。

与前面我们讲到的赝势函数不同，核实极化势 (CPP) 并不是一个简单的单电子算符，求解包含 CPP 的量子化学积分十分繁琐[34, 35]。由于缺少计算能量梯度的程序，目前 CPP 仅用于能量的单点计算。好在很多情况下我们并不需要 CPP，如当使用小核实赝势，即核实内电子数较少时。目前，对主族元素[36, 37]、镧系锕系元素[38] 以及第 11, 12 族元素[39, 40] 的大核实赝势，均有配套的 CPP 发表。对第 11, 12 族元素， Stoll 等还发展了包含电四极矩校正的 CPP[41]。

6.8　核实–核实 (核) 排斥势

对核实电子数很少的小核实赝势，核实–核实 (或核实–核) 间的重叠效应很小，相互作用 (V_{cc}) 可以直接采用点电荷模型求得 (见 (6.27) 式中的第一项)。对核实内电子数较多的大核实赝势，核实–核实相互作用仅用简单的 Coulomb 势表达就会产生较大的误差，此时必须考虑核实–核实 (或核实–核) 叠加效应 (又称彼此渗透效应)。为校正点电荷模型下的 Coulomb 排斥作用，引入核实–核实 (核) 排斥势$\Delta V_{cc}^{\lambda\mu}(r_{\lambda\mu})$，即 (6.27) 式中的第二项。

Stoll 等发现，将 $\Delta V_{cc}^{\lambda\mu}(r_{\lambda\mu})$ 写成简单的 Born-Mayer 势的形式即能有效地模拟成对的排斥校正[40, 39]

$$\Delta V_{cc}^{\lambda\mu}(r_{\lambda\mu}) = B_{\lambda\mu}\exp\left(-b_{\lambda\mu}r_{\lambda\mu}\right) \tag{6.64}$$

上式中的参数优化比较简单，只需将赝势得到的核实–核实 (或核实–核) 排斥能和真实的排斥能相比较，调整参数将差值最小化即可。

6.9　赝势的优化

在 6.5 节我们已经提到了由 Kahn, Baybutt, Truhlar 等提出的以波函数为优化目标的赝势构造方法。他们方法的缺点是得到的伪价电子波函数和真实的价电子波函数在价电子空间区域差别太大。为改进这一不足，后人提出了伪价电子波函数和真实波函数在价电子空间必须形状一致的概念，这就是我们将要讲到的从头算赝势的一个重要分支 —— 形状一致的从头算赝势。另外，1955 年 Preuss 提出了以能量为目标函数优化赝势的想法[42]，并于 1981 年发表了第一个基于能量优化 (称能量一致) 的半经验大核实赝势[43]。早期的能量一致赝势是以实验值作为参考值的，随着计算机技术的发展，此类半经验赝势已被以全电子从头算结果为参考值的从头算赝势代替。经过近 30 年的发展，能量一致的从头算赝势已成为赝势的另一个重要分支。

6.9.1　能量一致的从头算赝势

本节以 Stuttgart-Cologne 赝势为例，来说明能量一致从头算赝势参数优化的三个步骤。第一步是选择合适数目的核实电子和价电子。我们以铀 (U) 为代表来具体说明如何选择核实。表 6.1 列出了基态铀 ($5f^36d^17s^2$) 的 4s~7s 轨道的能量和轨道半径，表中没有列出能量更低的内层轨道 1s~3d 。

表 6.1　AE DHF 计算得到的铀基态 ($5f^36d^17s^2$) 轨道的能量 ϵ 和半径 $\langle r\rangle$

占据轨道①	ϵ/Hartree	$\langle r\rangle$/Bohr	占据轨道	ϵ/Hartree	$\langle r\rangle$/Bohr
4s	−54.32	0.32	5d−	−4.35	0.87
4p−	−48.19	0.32	5d	−4.04	0.90
4p	−39.54	0.35	5f−	−0.35	1.42
4d−	−29.73	0.35	5f	−0.32	1.46
4d	−28.13	0.36	6s	−2.14	1.50
4f−	−15.20	0.35	6p−	−1.34	1.66
4f	−14.79	0.36	6p	−0.98	1.90
5s	−12.60	0.67	6d−	−0.19	3.15
5p−	−10.13	0.70	6d	−0.18	3.29
5p	−8.09	0.78	7s	−0.20	4.34

注：1 Bohr = 5.291×10^{-11} m。

① 轨道符号的含义见表 2.2。

关于铀的核实选择有以下四种：

(1) 86 个核实电子的大核实。

从表 6.1 得知，5f, 6d, 7s 轨道的能量 (ϵ) 高于 −0.35 Hartree；1s~5d 轨道能量 (ϵ) 小于 −4.0 Hartree；6s, 6p 轨道能量 (ϵ) 小于 −1.3 Hartree。从能量的角度考虑，我们可以把能量明显低于 5f, 6d, 7s 等价层轨道的 1s~5d 轨道以及 6s, 6p 等共计 86 个内层轨道的电子都放入核实内，价层空间仅剩下 6 个价电子。

(2) 78 个核实电子的中等核实。

从表 6.1 我们发现，6s 和 6p 的轨道半径 $\langle r\rangle$ 分别为 1.50 Bohr, 1.66 Bohr, 1.90 Bohr，和 5f 的轨道半径 $\langle r\rangle$ (1.42 Bohr, 1.46 Bohr) 十分接近。考虑到 6s, 6p, 5f 轨道在空间区域的重叠效应，除价层轨道 5f, 6d, 7s 外，也应该把 6s 和 6p 轨道作价轨道处理，即价层空间包含 14 个价电子。

(3) 60 个核实电子的小核实。

由于内层轨道 5s~5d 和价电子轨道有较强的电子相关，精确计算也应该把 5s~5d 电子作价电子处理，这时的价电子数目达到 32 个。

(4) 78+n_{5f} 个核实电子的大核实，其中 n_{5f} 是 5f 轨道的电子布居数。

由于铀原子在参与化学成键时呈现不同的氧化态，在中等核实的基础上，我们还可以将不同氧化态时的 5f 轨道电子布居数放入核实中，此时的价电子数目为 $14-n_{5f}$。

接下来我们需要对不同核实可能带来的核实冻结误差进行评估。我们可以用全电子方法计算不同核实冻结条件下原子各组态的相对论能量，比较它们的差别并以此为标杆来估计赝势的核实冻结误差。表 6.2 列出了基于 Dirac-Colomb (DC) Hamilton 算符的多组态平均 Dirac-Hartree-Fock (MCDHF/DC) 计算得到的铀原子

和离子的不同电子态的相对能量以及不同核实冻结条件下的计算误差 (以下简称误差)[44]。相对其他轨道，容易发现 5f 轨道电子布居数的变化对计算结果影响最大，接下来我们就讨论误差随 5f 轨道电子布居数的变化。从节省计算量的角度，当然希望核实越大越好。从表 6.2 可以看出，对所有包含 $5f^3$ 的电子组态，仅含 6 个价电子的大核实的误差最高为 0.35 eV，包含 14 个价电子的中等核实的误差小于 0.001 eV，与含 32 个价电子的小核实的误差基本一致。

表 6.2　MCDHF/DC 得到的铀的各组态相对于基态 (U[Rn] $5f^3 6d^1 7s^2$) 的能量 (eV)，以及大 ($Q=6$，$Q=14-n_{5f}$)、中 ($Q=14$)、小 ($Q=32$) 三种核实的核实冻结误差 (eV)

电荷	组态			AE DHF	核实冻结误差			
	5f	6d	7s	DC	$Q=32$	$Q=14$	$Q=6$	$Q=14-n_{5f}$
6+				167.9498	0.0077	1.2479	15.4283	0.0059
5+		1		118.1617	0.0058	1.0291	11.3885	0.0129
4+		2		78.6962	0.0048	0.8850	8.2895	0.0040
3+		3		49.0026	0.0043	0.7963	6.0066	0.0013
2+		4		28.4542	0.0040	0.7488	4.4237	0.0010
1+		4	1	16.0181	0.0039	0.7379	3.8610	0.0002
0		4	2	9.7540	0.0038	0.7365	3.5686	0.0000
5+	1			107.4044	0.0021	0.4545	7.1973	0.1523
4+	1	1		69.4125	0.0014	0.3779	4.8002	0.0626
3+	1	2		41.0364	0.0014	0.3359	3.1180	0.0175
2+	1	3		21.6260	0.0013	0.3179	2.0282	0.0023
1+	1	3	1	9.6394	0.0013	0.3082	1.6778	0.0003
0	1	3	2	3.6701	0.0013	0.3046	1.5126	0.0000
4+	2			62.9423	0.0003	0.0886	2.3845	0.1796
3+	2	1		35.8785	0.0002	0.0765	1.2558	0.0535
2+	2	1	1	17.5395	0.0002	0.0682	0.8277	0.0260
1+	2	1	2	5.7548	0.0001	0.0636	0.5841	0.0126
0	2	2	2	0.4682	0.0002	0.0678	0.3493	0.0000
3+	3			33.3768	0.0000	0.0006	0.3479	0.1458
2+	3		1	15.7154	0.0000	0.0006	0.1366	0.0698
1+	3		2	4.4832	0.0000	0.0009	0.0412	0.0306
0	3	1	2	0.0000	0.0000	0.0000	0.0000	0.0000
2+	4			17.2536	0.0002	0.0467	0.2082	−0.0399
1+	4		1	6.8681	0.0001	0.0450	0.2224	−0.0193
0	4		2	2.0435	0.0002	0.0426	0.2430	0.0000
1+	5			12.5023	0.0004	0.0945	0.6651	0.1921
0	5		1	8.1022	0.0004	0.0761	0.5418	0.0000

注：MCDHF/DC 指基于 Dirac-Coulomb Hamilton 算符的多组态 DHF 计算。

冻结的核实：U $Q=6$，1s~5d, 6s, 6p; $Q=14$，1s~5d; $Q=32$，1s~4d; $Q=14-n_{5f}$，1s~5f。n_{5f} 为 5f 轨道电子布居数。

显然，如果我们只考虑铀 (U) 的 5f 轨道电子布居数保持基态 ($5f^3$) 的情形，中等核实，甚至大核实也可以取得令人满意的结果。众所周知的是，在化学反应中铀以及其他轻的锕系元素的 5f 轨道电子布居数常发生改变，所以我们还要检查 5f 轨道电子布居数非基态占据时不同核实的误差。从表 6.2 我们得知，如果 5f 轨道电子布居数降为 2 或升为 4，此时小核实的误差小于 0.0002 eV，中等核实误差小于 0.1 eV，二者均能满足一般化学研究要求的精度。如果需要计算结果精度大于 0.1 eV，则需要采用小核实。此时大核实的误差明显增大，对正三价和正四价的铀 (U^{3+}, U^{4+})，误差可达到 1.3 eV 和 2.4 eV，这意味着仅含 6 个价电子的大核实已失去使用价值。

以上结论也可以推广到 $5f^1$, $5f^0$, $5f^5$ 等其他情况。从表 6.2 可以看出，无论对应何种电子组态小核实的误差始终低于 0.008 eV，而其他两种核实的误差则显著增大。对化学反应中常见的 U^{5+} 和 U^{6+}，中等核实的误差达到了 1.03 eV 和 1.25 eV，大核实的误差则更大。由此我们得出结论：只有小核实能精确描述铀的所有氧化态。

以上讨论的是核实冻结误差随 5f 轨道电子布居数的变化。在选择相同的 5f 轨道电子布居数的条件下，如果计算更多包含不同 6d 轨道电子数的电子组态 (表 6.2 中未列出)，可以发现核实冻结误差有轻微不同。在表 6.2 中我们也能看到 6d 轨道电子布居数对核实冻结误差的影响，总的来说 6d 轨道对核实冻结误差的影响要远小于 5f 轨道的影响，而且这种影响也不妨碍我们已得到的结论。

当研究体系很大时，将 5f 电子作价电子处理的小核实带来的计算量仍然是惊人的，为此我们可以选择另一种价电子数目为 $14 - n_{5f}$ 的大核实，即将 5f 轨道放入核实中，n_{5f} 代表 5f 轨道电子布居数。从表 6.2 可以看出，对所有参考态，大核实的核实冻结误差最大为 0.2 eV。对 $5f^1$ 或 $5f^0$ 的所有电子组态，它的误差要小于中等核实。随着 5f 轨道电子布居数的增加，核实误差也随之增大。关于核实冻结误差的更深入讨论，读者可参见文献 [5] 中的 247~252 页。

能量一致赝势优化的第二步为选择合适的参考态和计算参考态能量。参考态的选择是和赝势的使用范围密切相关的，一般规律是要包括该元素的中性原子、阴离子 (如果存在的话)、低价的阳离子等。还是以铀为例，铀的小核实赝势主要用于小分子的精确计算，所选参考态范围也最广，包括了铀的中性原子 (U)、正一价到正七价阳离子 (U^{1+}, U^{2+}, U^{3+}, U^{4+}, U^{5+}, U^{6+}, U^{7+}) 的基态、激发态等总计 100 个非相对论电子组态和 30 190 个相对论 J 能级[45]。含 10 个价电子的大核实赝势仅适用于体系中铀为正二价的情形，所选参考态数目也远小于小核实赝势，即 U, U^{1+}, U^{2+} 的 9 个非相对论组态[46]。对适用于体系中铀为正三价的大核实赝势 (含 11 个价电子)，所选参考态则包含 U, U^{1+}, U^{2+}, U^{3+} 的基态、激发态等总计 14 个非相对论组态[47]。含 12 个价电子的大核实赝势适用于铀的氧化态为正四价的情

形，所选参考态包括了 U, U^{1+}, U^{2+}, U^{3+}, U^{4+} 的 18 个非相对论组态[46]。

参考态能量的计算也和赝势的使用密切相关。以 U 为例，由于我们希望小核实赝势可用于二分量 (即包含旋–轨耦合) 的变分计算，所以我们选择基于 DCB(Dirac-Coulomb Hamilton 加上 Breit 校正)Hamilton 算符的 MCDHF(多组态的 Dirac-Hartree-Fock) 方法来计算参考态能量[45]。为避免基组对参考态能量的影响，我们选择了采用数值解法的四分量的有限差分 MCDHF 程序 GRASP[48]。

我们也可以将二分量的赝势写成自旋平均的一分量赝势和旋–轨耦合赝势两项 (见 6.6 节) 再分别优化。对一分量赝势我们选择的参考态能量来自基于 Wood-Boring (WB) Hamilton 算符[49] 的相对论 HF 计算 (不包括旋–轨耦合)。旋–轨耦合赝势的参考态能量则来自 WB Hamilton 算符的旋–轨耦合算符[47, 46, 50]。与二分量赝势优化一致，为避免基组的影响，我们选择有限差分 HF 程序 MCHF77[51] 得到参考态能量。除 WB Hamilton 算符以外，原则上也可以选择其他形式的全电子 Hamilton 算符，如 DKH Hamilton 算符[52] 来计算参考态能量，但目前并没有数值求解 DKH Hamilton 算符的有限差分程序。

在得到全电子参考态能量后就可以进行能量一致赝势优化的最后一步，即赝势参数的优化。对每一个参考态 I，假定 E_I^{PP} 代表由初始猜测赝势参数得到的能量，E_I^{AE} 代表全电子计算得到的能量，我们对每一个参考态的均方差求和，得

$$S = \sum_I w_I (E_I^{\mathrm{PP}} - E_I^{\mathrm{AE}} + \Delta E_{\mathrm{shift}})^2 = \min \tag{6.65}$$

式中，$\Delta E_{\mathrm{shift}}$ 是整体平移参数，通常是基态价电子能量的 1% 左右，用于提高优化的精度。优化赝势参数求 (6.65) 式极小值的过程可以用最小二乘法等数学上的标准方法来解决。参数优化的最后精度取决于核实的大小和我们采用的半局域赝势函数的精度，对多电子组态而言，赝势精度一般为 0.1 eV 左右，而对单电子态，赝势的能量误差小于 0.001 eV。

6.9.2 形状一致的从头算赝势

与能量一致赝势以全电子价电子组态能量为优化目标不同，形状一致的从头算赝势以全电子价电子轨道和轨道能量为参考值[19, 18]，这一点我们在 6.5 节中已经提到。形状一致的从头算赝势不仅要求伪价电子轨道 (φ_{p}) 的能量 (ϵ_{p}) 和全电子的价电子轨道 (φ_{v}) 的能量 (ϵ_{v}) 相同，而且要求在核实外的区域 ($r > r_{\mathrm{c}}$)，φ_{p} 和 φ_{v} 保持形状一致。具体来说，对一给定的 lj 态 (l, j 分别代表角量子数和总量子数，取值范围见 6.6 节)

$$\varphi_{\mathrm{p},lj}(r) = \begin{cases} \varphi_{\mathrm{v},lj}(r) & (r \geqslant r_{\mathrm{c}}) \\ f_{lj}(r) & (r < r_{\mathrm{c}}) \end{cases} \quad 且 \quad \epsilon_{\mathrm{p},lj} = \epsilon_{\mathrm{v},lj} \tag{6.66}$$

辅助函数 f_{lj} 在核实区域 $(r < r_c)$ 是光滑、无节点的。f_{lj} 函数形式和 r_c 的选取通常依靠经验，稍后我们会具体举例说明。

对给定的 lj，形状一致的半局域赝势 $V_{lj}(r)$ (见 (6.38) 式) 通过求解径向的 Fock 方程得到

$$\left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{l(l+1)}{2r^2}+V_{lj}^{\mathrm{PP}}(r)+\hat{W}_{\mathrm{p},lj}(\{\varphi_{\mathrm{p}',l'j'}\})\right]|\varphi_{\mathrm{p},lj}(r)\rangle=\epsilon_{\mathrm{v},lj}|\varphi_{\mathrm{p},lj}(r)\rangle \quad (6.67)$$

上式中的前两项为径向的动能算符，$\hat{W}_{\mathrm{p},lj}$ 代表由 $\varphi_{\mathrm{p},lj}$ 表达的有效价电子 Coulomb 势和交换势。由于全电子参考态计算已经严格考虑了相对论效应，V_{lj}^{PP} 也隐含了相对论效应的贡献。为避免基组对计算结果的影响，一般采用数值法求得全电子价层轨道 $(\varphi_{\mathrm{v},lj})$，所以由 (6.67) 式得到的伪价层轨道 (V_{lj}^{PP}) 也是以分立的数值点来表达的。

重复以上过程直到对所有的 lj 都求得 V_{lj}^{PP}，其中 $l=0,1,\cdots,L-1; j=\left|l-\frac{1}{2}\right|,\left|l+\frac{1}{2}\right|$，$(L-1)$ 是核实内电子轨道的角量子数最大值。下一步就是将 V_{lj}^{PP} 拟合成 Gauss 函数的线性组合

$$\hat{\hat{V}}^{\mathrm{PP}}(\boldsymbol{r}_i)=-\frac{Q}{r}+\sum_{lj}\left(\sum_k A_{lj,k}r^{n_{lj,k}-2}\mathrm{e}^{-\alpha_{lj,k}r^2}\right)\hat{P}_{lj} \quad (6.68)$$

如果希望得到自旋平均的一分量赝势，只需将 (6.67) 式和 (6.68) 式中所有变量的脚标由 lj 改为 l(见 6.5 节)，再重复与求二分量赝势的相同步骤，就得到了半局域一分量赝势函数 $\Delta\hat{V}_{\mathrm{cv,av}}$ (见 (6.43) 式)。

20 世纪 80~90 年代，Christiansen, Ermler 等陆续发表了基于 DC/DHF 全电子单组态计算的形状一致的赝势及相应的旋–轨耦合算符[53−62]。同一时期，对主族元素和过渡金属元素 Hay 和 Wadt 发表了基于 Cowan-Griffin (CG) Hamilton 算符[21] 自旋平均计算的形状一致的赝势[63−65]，这也是目前最广泛使用的赝势之一，在基组数据库中通常采用缩写 LANLX (Los Alamos National Laboratory, X=1, 2; 1 代表大核实，2 代表小核实)。2008 年，他们又重新优化得到了特别适用于 DFT 计算的基组[66]。

我们以 Hay 和 Wadt 的形状一致赝势为例，介绍 (6.66) 式中 r_c 和辅助函数 f_l 的选取。

对每一角量子数 l，先求得轨道 φ_l 的径向最大值，与此对应的 $r_{\max}$ 定义为 r_c。当 $r < r_c$ 时，辅助函数 f_l 定义为 r 的多项式

$$f_l(r)=r^b\left(a_0+a_1r+a_2r^2+a_3r^3+a_4r^4\right) \quad (6.69)$$

在非相对论级别，上式中 $b=l+3$；在相对论级别，当 $l\neq 0$ 时 $b=\lambda+2$；当 $l=0$

时，即相对论的 s 轨道，$b=\lambda+3$。参数 λ 由下式决定

$$\lambda+1=\frac{1}{2}(1-\delta_{0,l})+\sqrt{l(l+1)+\frac{1}{4}(1+\delta_{l,0})^2+(\alpha Z)^2} \tag{6.70}$$

式中，α 是精细结构常数。函数 f_l 中的五个系数 a_i 按以下条件确定

(1) $\varphi_{\rm p}(r)$ 必须是归一化的。

(2) 当 $r=r_{\rm c}$ 时，$f_l(r)=\varphi_{\rm v}(r)$，且二者对 r 的一、二、三阶导数也相等。

Hay 和 Wadt 将伪价层轨道拟合为 Gauss 函数线性组合以得到更为紧凑的赝势价电子基组

$$\varphi_{{\rm p},l}=\sum_i C_i r^l {\rm e}^{-\alpha_i r^2} \tag{6.71}$$

为使赝势 ($\hat{\tilde{V}}^{\rm PP}$) 的 Gauss 函数展开形式 (见 (6.68) 式) 更简单，Toulouse 研究小组引入了算符 $\hat{\mathcal{O}}$，优化目标则是最小化 $\hat{\mathcal{O}}$ 的模[67]

$$\|\hat{\mathcal{O}}\|=\langle\varphi_{{\rm p},lj}|\hat{\mathcal{O}}^2|\varphi_{{\rm p},lj}\rangle^{1/2} \tag{6.72}$$

其中

$$\hat{\mathcal{O}}=\tilde{\epsilon}_{{\rm v},lj}|\tilde{\varphi}_{{\rm p},lj}\rangle\langle\tilde{\varphi}_{{\rm p},lj}|-\epsilon_{{\rm v},lj}|\varphi_{{\rm p},lj}\rangle\langle\varphi_{{\rm p},lj}| \tag{6.73}$$

式中，$\varphi_{{\rm p},lj}$ 和 $\epsilon_{{\rm v},lj}$ 代表由真实赝势 $V_{lj}^{\rm PP}$ 得到的波函数和能量。所谓真实赝势 $V_{lj}^{\rm PP}$ 就是求解径向 Fock 方程 (6.67) 式得到的数值赝势 $V_{lj}^{\rm PP}$。$\tilde{\varphi}_{{\rm p},lj}$ 和 $\tilde{\epsilon}_{{\rm v},lj}$ 代表由赝势分析表达式 $\tilde{V}_{lj}^{\rm PP}$(见 (6.68) 式) 得到的波函数和能量。按这种方法优化，对每一个 lj 最多只需要 3 个展开式就可以了。优化得到的 $\tilde{\varphi}_{{\rm p},lj}$ 与 $\varphi_{{\rm p},lj}$ 的内积 $\langle\tilde{\varphi}_{{\rm p},lj}|\varphi_{{\rm p},lj}\rangle\approx 1$(误差小于 10^{-5})，本征值差别 $(\tilde{\epsilon}_{{\rm v},lj}-\epsilon_{{\rm v},lj})$ 小于 10^{-3} Hartree。采用 Toulouse 研究小组的优化方案，Stevens 等发表了基于全电子 DC/DHF 计算的形状一致的赝势[68−70]。

6.9.3　广义的相对论的有效核实势

对过渡金属、镧系、锕系等元素，当选择小核实赝势时，在核实外区域可能出现所选电子组态波函数 ($\varphi_{{\rm v},lj}$) 存在径向节点的情况，此时由径向 Fock 方程 (6.67) 式反演求解赝势 $V_{lj}^{\rm PP}$ 就会出现奇点。通常的解决方案是避开 $\varphi_{{\rm v},lj}$ 中存在节点的参考态，如几乎所有的形状一致的赝势都只选择元素的正离子态为参考态。但是，仅选择正离子参考态得到的赝势在用于中性原子或分子计算时有可能产生较大的核实冻结误差。另外，前面我们讲到的所有赝势投影算符中只包含了量子数 lm (或 ljm)，并不包含主量子数 n，也就是说赝势算符作用到核实外 n 不同而 lm (或 ljm) 相同的轨道的效果是一样的。严格地说，这只是一种近似，在靠近核实 $(r\approx r_{\rm c})$ 附

近，对 n 不同而 lm (或 ljm) 相同的轨道，由径向 Fock 方程 (6.67) 式反演求得的赝势会有细微差别[71]。

为解决以上问题，Titov, Mosyagin 等提出了广义的相对论的有效核实势 (generalized relativistic ECP，GRECP)[71−74] 方法，其特点是

(1) 在径向 Fock 方程奇点附近引入赝势的展开式。

(2) 将核实外电子细分为核实外层电子、价层电子两部分，作用到二者的赝势算符并不相同。

用于原子计算的 GRECP 算符可写成[71]

$$\begin{aligned}\Delta\hat{V}_{\text{cv}}(\boldsymbol{r}_i) \cong V_{n_\text{v}LJ} &+ \sum_{l=0}^{L-1}\sum_{j=|l-\frac{1}{2}|}^{l+\frac{1}{2}}[V_{n_\text{v}lj}(r_i) - V_{n_\text{v}LJ}(r_i)]\,\hat{P}_{lj}\\ &+ \sum_{n_\text{oc}}\sum_{l=0}^{L-1}\sum_{j=|l-\frac{1}{2}|}^{l+\frac{1}{2}}[V_{n_\text{oc}lj}(r_i) - V_{n_\text{v}lj}(r_i)]\,\hat{\tilde{P}}_{n_\text{oc}lj}\end{aligned} \tag{6.74}$$

式中，n_v 和 n_oc 分别代表价层和核实外层电子轨道的主量子数。$J = L + 1/2$, 投影算符 $\hat{P}_{lj}$, $\hat{\tilde{P}}_{n_\text{oc},lj}$ 分别作用到二分量的自旋本征函数 χ_{ljm} 以及核实外层伪自旋本征函数 $\tilde{\phi}_{n_\text{oc},lj}$

$$\hat{P}_{lj} = \sum_{m=-j}^{j}|ljm\rangle\langle ljm| \tag{6.75}$$

$$\hat{\tilde{P}}_{n_\text{oc},lj} = \sum_{m=-j}^{j}|\widetilde{n_\text{oc}, ljm}\rangle\langle\widetilde{n_\text{oc}, ljm}| \tag{6.76}$$

由 (6.74) 定义的赝势算符并不是一个厄米算符，还需近一步近似得到厄米的赝势算符[71]

$$\begin{aligned}V^{\text{GRECP}} = V_{n_\text{v}LJ} + \sum_{l=0}^{L}\sum_{j=|l-1/2|}^{l+1/2}\Bigg\{&\Big[V_{n_\text{v}lj}(r) - V_{n_\text{v}LJ}(r)\Big]\hat{P}_{lj}\\ &+\sum_{n_\text{oc}}\Big[V_{n_\text{oc},lj}(r) - V_{n_\text{v},lj}(r)\Big]\hat{\tilde{P}}_{n_\text{oc},lj}\\ &+\hat{\tilde{P}}_{n_\text{oc},lj}\sum_{n_\text{oc}}\Big[V_{n_\text{oc},lj}(r) - V_{n_\text{v},lj}(r)\Big]\\ &-\sum_{n_\text{oc},n'_\text{oc}}\hat{\tilde{P}}_{n_\text{oc},lj}\left[\frac{V_{n_\text{oc},lj} + V_{n'_\text{oc},lj}}{2} - V_{n_\text{v},lj}(r)\right]\hat{\tilde{P}}_{n'_\text{c},lj}\Bigg\}\end{aligned} \tag{6.77}$$

将 (6.77) 式与标准的半局域赝势表达式 (6.40) 式相比，GRECP 方法增加了三个局域赝势表达项 (即 (6.77) 式中的第 2~4 行)，这三项考虑了量子数 lj 相同时作用

到核实外层电子和价电子的有效核实势的区别。用于分子计算的 GRECP 算符通常写成含自旋平均和旋–轨耦合两项的和[71]。GRECP 的优化方法和形状一致的从头算赝势的优化方法类似[71]，在此就不再赘述。需要指出的是，由于 GRECP 包括了奇点及对核实外层电子和价电子的区别处理，GRECP 得到的赝势曲线在核实 ($r \approx r_{\mathrm{c}}$) 附近是不光滑的，这就给随后的参数化赝势 (即 Gauss 函数展开) 带来了困难[71]。为保证拟合精度，GRECP 需要更多的 Gauss 函数，如对铀元素，自旋平均的 GRECP 参数的数目达到了 158 个，而相同精度级别的能量一致的赝势仅需要 52 个参数。

对 U, Pu 以及 E112, E113, E114 等超重元素，Titov 和 Mosyagin 等已发表了包括 Breit 相互作用的 GRECP[75, 76]。由于目前大多数的标准量子化学程序并不支持 GRECP 方法，所以该方法的使用并不普遍[75,77−81]。

6.10　价电子基组

赝势方法只对核实外价电子作严格的量子化学处理，也就是说我们只需为价电子计算提供基组，所以与赝势配套的基组又称价电子基组。既然需要精确处理的电子数目减少了，价电子基组的规模也就自然要小于全电子基组。价电子基组函数形式的选取则与全电子基组类似，目前几乎所有赝势的价电子基组都采用 Gauss 函数形式。Gauss 函数在直角坐标系中的表达形式是

$$g(\alpha, l, m, n; x, y, z) = N\mathrm{e}^{-\alpha r^2} x^l y^m z^n \tag{6.78}$$

式中，N 是归一化常数；α 称为指数；x，y，z 代表坐标；l，m，n 为整数；$r^2 = x^2 + y^2 + z^2$。用于赝势基组表达的 Gauss 函数也称为精确 Gauss 函数，通常从原子的 HF 赝势计算获得，即优化指数以最小化参考态的能量。

精确的 Gauss 函数基组一般仅用于原子的电子结构计算，用于分子电子结构计算的基函数是精确 Gauss 函数的线性组合，即收缩的 Gauss 基组。线性组合系数多来自对所选原子组态态密度矩阵的平均 (如 Stuttgart-Cologne 从头算能量一致赝势)。基组的收缩方案有两种，即“一般 (general) 收缩”和“分段 (segmented) 收缩”，收缩方案的选择与基组的用途相关。对前者，每一个精确 Gauss 函数都出现在收缩后的每一个基函数中；对后者，每一个精确 Gauss 函数一般只在收缩后的基函数中出现一次。与后者相比，前者可以选择原子的自然轨道 (atomic natural orbitals，ANO)，得到的收缩基组能很好地描述分子的相关能[82−84]。遗憾的是，现有量子化学软件中的积分程序普遍对后者更有效，所以在积分快慢决定计算耗时的 HF 和 DFT 计算中，人们常使用分段收缩的基组。

价电子基组的优化和赝势优化同等重要，所以一般赝势和配套基组是同时发

表的。在本章一开始，我们已提到伪价电子轨道的节点数目要少于真实的价电子轨道，所以相同计算精度下不仅价电子基组的数目要小于全电子基组，基组重叠误差①(basis set superposition error，BSSE) 也较小。

图 6.2 比较了赝势及全电子计算得到的铅 (Pb) 基态 ($6s^2 6p^6$) 的 6s 轨道。从图中可以看出，在核实外，无论是能量一致赝势 (Stuttgart-Cologne)，还是形状一致赝势 (LANL)，计算得到的伪 6s 轨道和全电子结果符合得相当好，但在核实内区域，能量一致赝势轨道和形状一致赝势轨道明显不同，这就是说，如果将为某一特定赝势优化的基组用于其他赝势时，即使核实的定义相同，也可能产生较大的误差，所以赝势及配套基组必须同时使用。

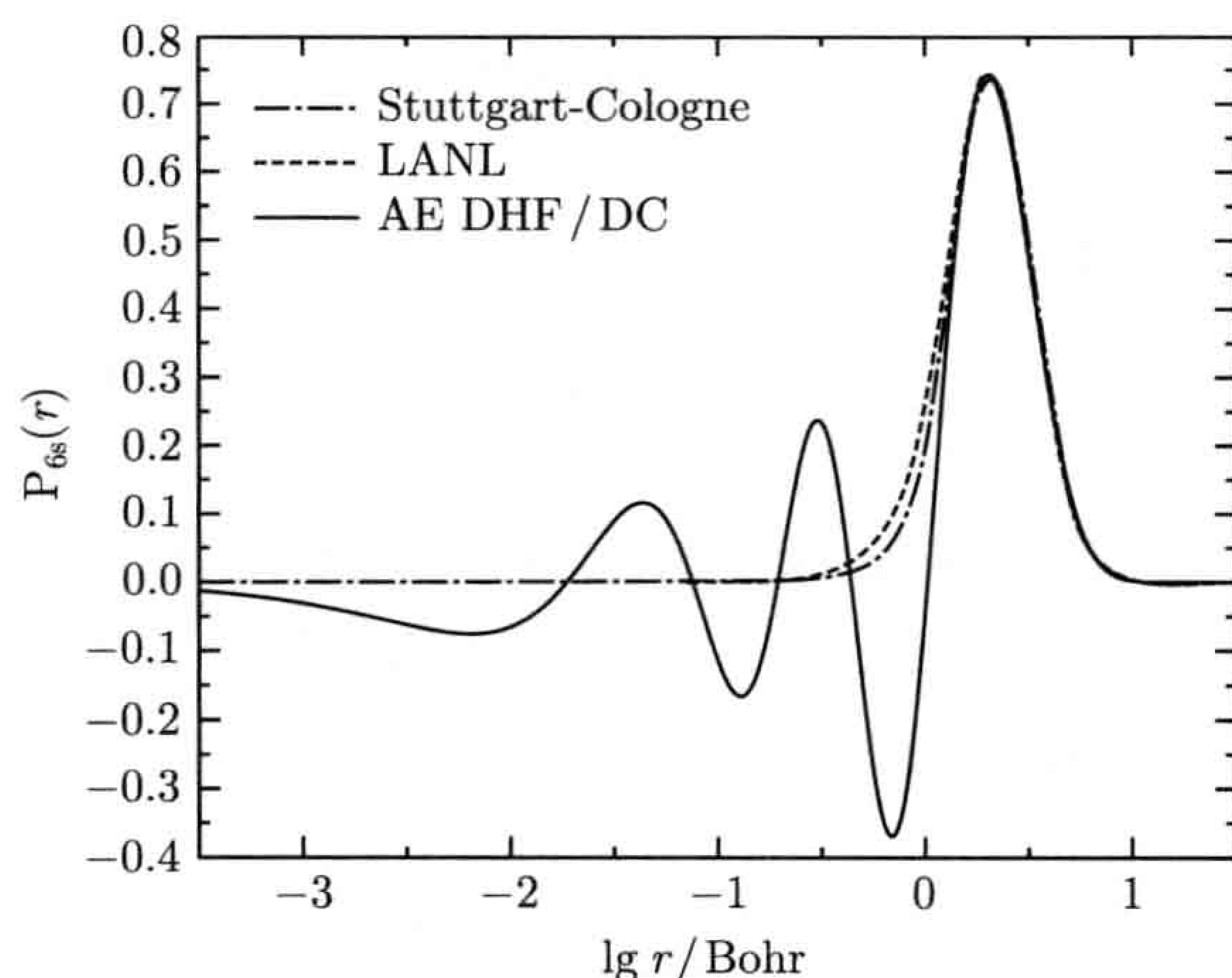


图 6.2 能量一致赝势 (Stuttgart-Cologne)、形状一致赝势 (LANL)、全电子计算 (AE DHF/DC) 得到的 Pb 基态 ($6s^2\ 6p^2$) 6s 轨道径向函数 $P_{6s}(r)$

对一些广泛使用的赝势，为满足不同计算级别的要求，人们也优化了不同精度的价电子基组。例如，对周期表中除镧系元素以外的从 37 号元素 Rb 到 86 号元素 Rn 的能量一致赝势，Weigend 和 Ahlrichs 优化了分裂价基、三 -Zeta、四 -Zeta 等三个不同精度的分段收缩基组[85, 86]；对从 Ga 到 Kr 的能量一致赝势，Sundermann 和 Martin 发表了相关能一致的一般收缩基组[87]；对所有现代的能量一致小核实赝势，Peterson 等优化得到了不同精度水平的相关能一致的一般收缩基组[88, 89]；对形状一致赝势，Hay 等优化了适用于 DFT 计算的基组[66]；等等。

随着互联网的普及，赝势和基组的获得也变得十分方便。下面列出了几个常用的赝势 (基组) 网站，供读者免费下载。

① 以双原子分子 AB 为例，假设我们使用有限基组优化计算原子间距离 R。显然，当 R 较大时，A 和 B 的基组重叠积分要小于 R 较小时的基组重叠积分，所以优化得到的平衡键长 R_e 将偏短，这种由于使用有限基组而导致的原子间距离的人为改变称为基组重叠误差

(1) Stuttgart-Cologne 赝势和基组

http://www.tc.uni-koeln.de/PP/index.en.html

(2) Peterson 相关能一致基组

http://tyr0.chem.wsu.edu/ kipeters/basis.html

(3) Weigend 分段收缩基组

http://www.cosmologic.de/basis-sets/basissets.php

(4) 基组交换网站

http://bse.pnl.gov/bse/portal

6.11 模型势方法

6.11.1 Huzinaga-Cantu 方程

除了赝势 (PPs) 方法以外，模型势 (model potentials，MPs) 方法也是有效核实势 (ECPs) 方法的一个重要分支。MPs 的概念最早是由 Bonifacic 和 Huzinaga 于 1974 年提出[90]，它的理论基础则源于 Huzinaga-Cantu (HC) 方程[91, 92]。该方程的推导和我们在 6.1 节讲过的 Phillips-Kleinman (PK) 方程 (6.8) 式的推导十分类似，这里不再重复，仅给出 HC 方程形式

$$\left[\hat{H}_{\text{eff}}+\sum_{\text{c}}(-2\epsilon_{\text{c}})|\varphi_{\text{c}}\rangle\langle\varphi_{\text{c}}|\right]|\varphi_{\text{v}}\rangle=\epsilon_{\text{v}}|\varphi_{\text{v}}\rangle \tag{6.79}$$

式中，$\hat{H}_{\text{eff}}$ 代表一个有效的单电子 Hamilton 算符，如可以是 Fock 算符。价电子本征函数 ($|\varphi_{\text{v}}\rangle$) 和核实电子本征函数 ($|\varphi_{\text{c}}\rangle$) 满足 (6.1) 式和 (6.2) 式。

HC 方程告诉我们，将 $\hat{H}_{\text{eff}}$ 增加一项可得一个新的算符，原算符的本征函数 $|\varphi_{\text{v}}\rangle$ 仍是新算符的本征函数，本征值不变。如果将新算符作用到 $|\varphi_{\text{c}}\rangle$，不难发现 $|\varphi_{\text{c}}\rangle$ 也是该算符的本征函数，但本征值不再是 ϵ_{c}，而是 $-\epsilon_{\text{c}}$ 。所以新算符中的第二项又称为能量位移算符，其作用是将核实电子轨道 ($|\varphi_{\text{c}}\rangle$) 平移到到能量为 $-\epsilon_{\text{c}}$，即正能量的区域。由于所有价电子轨道能量都小于 0，变分计算时就能保证价电子不会塌缩成核实电子。

HC 方程和 Phillips-Kleinman (PK) 方程 (6.8) 式十分形似，都是在真实的有效电子 Hamilton 算符后增加一项，得到模型 Hamilton 算符。对价电子轨道而言，该模型 Hamilton 算符和原 Hamilton 算符的本征值相同。比较两个方程我们发现，HC 方程中模型 Hamilton 的价电子本征函数为真实的价电子波函数 ($|\varphi_{\text{v}}\rangle$)，而 PK 方程的价电子本征解为伪价电子波函数 ($|\varphi_{\text{p}}\rangle$)。由于 $|\varphi_{\text{p}}\rangle$ 不必保留 $|\varphi_{\text{v}}\rangle$ 的节点结构，所以与 MPs 配套的价电子基组通常要比赝势的价电子基组大一些。

6.11.2 多电子体系的价电子模型 Hamilton 算符

假定一个原子 (核电荷数为 Z)，由 n_v 个价电子和包含 n_c 个电子的核实组成，核实所带电荷为 Q，$Q = Z - n_\text{v}$。在固定核实近似下，该原子的价电子模型 Hamilton 算符可写为

$$\hat{H}_\text{v} = -\frac{1}{2}\sum_i^{n_\text{v}} \nabla_i^2 + \sum_{i<j}^{n_\text{v}} \hat{g}(i,j) - \sum_i^{n_\text{v}} \frac{Q}{r_i} + \sum_i^{n_\text{v}} \left[\Delta\hat{V}_\text{C}(i) + \Delta\hat{V}_\text{X}(i)\right] \tag{6.80}$$

式中，$\Delta\hat{V}_\text{C}(i)$ 代表核实电子和价电子的 Coulomb 相互作用

$$\Delta\hat{V}_\text{C}(i) = -\frac{n_\text{c}}{r_i} + 2\sum_\text{c} \hat{J}_\text{c}(i) \tag{6.81}$$

$\Delta\hat{V}_\text{X}(i)$ 描述核实电子和价电子的交换 (exchange) 能

$$\Delta\hat{V}_\text{X}(i) = -\sum_\text{c} \hat{K}_\text{c}(i) \tag{6.82}$$

算符 $\hat{J}_\text{c}(i)$ 和 $\hat{K}_\text{c}(i)$ 表示与核实电子轨道 $|\varphi_\text{c}\rangle$ 相联系的标准 Coulomb 算符和交换算符。由 HC 方程 (6.79) 式，将投影 (位移) 算符 $\hat{P}(i)$ 引入 $\hat{H}_\text{v}$ (见 (6.80) 式)

$$\hat{P}(i) = \sum_\text{c} (-2\epsilon_\text{c})|\varphi_\text{c}(i)\rangle\langle\varphi_\text{c}(i)| \tag{6.83}$$

对电子 i，定义模型势 $\hat{V}^\text{MP}(i)$

$$\hat{V}^\text{MP}(i) = \Delta\hat{V}_\text{C}(i) + \Delta\hat{V}_\text{X}(i) + \hat{P}(i) \tag{6.84}$$

我们就得到了原子的价电子模型 Hamilton 算符

$$\hat{H}_\text{v}{}^\text{MP} = -\frac{1}{2}\sum_i^{n_\text{v}} \nabla_i^2 + \sum_{i<j}^{n_\text{v}} \hat{g}(i,j) - \sum_i^{n_\text{v}} \frac{Q}{r_i} + \sum_i^{n_\text{v}} \hat{V}^\text{MP}(i) \tag{6.85}$$

分子的模型势定义与分子的赝势定义类似 (见 6.4 节)，即在核实冻结近似条件下，将分子的模型势写成是所有原子模型势的和。分子的价电子模型 Hamilton 算符可写为[1]

$$\hat{H}_\text{v}^\text{MP} = -\frac{1}{2}\sum_i^{n_\text{v}} \nabla_i^2 + \sum_{i<j}^{n_\text{v}} \hat{g}(i,j) + \sum_{\lambda<\mu} \frac{Q_\lambda Q_\mu}{r_{\lambda\mu}} + \sum_i \sum_\lambda \left[-\frac{Q^\lambda}{r_{\lambda i}} + \hat{V}^{\text{MP},\lambda}(i)\right] \tag{6.86}$$

式中，$\hat{V}^{\text{MP},\lambda}(i)$ 代表核实 λ 的模型势。6.11.3 节我们介绍 (6.84) 式中各项的具体表达。

6.11.3 从头算模型势

模型势的分析表达式存在三种不同的形式，由此产生了模型势的不同分支[93−96]，如模型核实势 (model core potentials，MCPs)、从头算模型势 (*ab initio* model potentials，AIMPs) 等。关于它们的详细介绍读者可以阅读 Klobukowski, Huzinaga, Sakai[1]，以及 Seijo, Barandiarán 等的综述文章[2]。本节我们以应用最广泛的 AIMPs 为例，介绍模型势的解析形式。

AIMPs 方法在 1987 年由 Huzinaga, Seijo, Barandiarán 等提出[93]。他们建议用一个局域的球对称模型势来表达核实电子和价电子的 Coulomb 相互作用，即 (6.84) 式中的 $\Delta\hat{V}_{\mathrm{C}}^{\lambda}(\boldsymbol{r}_i)$

$$\Delta\hat{V}_{\mathrm{C}}^{\lambda}(\boldsymbol{r}_i)=\Delta V_{\mathrm{C}}^{\lambda}(r_{\lambda i})=\frac{1}{r_{\lambda i}}\sum_k C_k^{\lambda}\mathrm{e}^{-\alpha_k^{\lambda}r_{\lambda i}^2} \tag{6.87}$$

上式中的指数 α_k^{λ} 和系数 C_k^{λ} 根据全电子结果优化，且满足 $\sum\limits_k C_k^{\lambda}=Z_{\lambda}-Q_{\lambda}=n_{\mathrm{c}}^{\lambda}$($n_{\mathrm{c}}^{\lambda}$ 为第 λ 个核实内的电子数目)。上述局域势的积分计算并不复杂，可以很容易通过增加展开式的数目 k 以提高优化精度。

(6.84) 式中的第二项，即非局域的交换势 $\Delta\hat{V}_{\mathrm{X}}^{\lambda}(i)$ 由一组定义在核实 λ 的函数 χ_p^{λ} 表达

$$\Delta\hat{V}_{\mathrm{X}}^{\lambda}(\boldsymbol{r}_i)=\sum_{p,q}|\chi_p^{\lambda}(i)\rangle A_{pq}^{\lambda}\langle\chi_q^{\lambda}(i)| \tag{6.88}$$

由于交换能的短程性，展开函数 χ_p^{λ} 的数目不用太大。展开函数 χ_p^{λ} 通常取为价电子基组的精确 Gauss 函数，此时核实交换算符产生的单电子积分和真实的全电子核实交换算符是一样的，双中心和三中心积分则必须作近似处理。A_{pq}^{λ} 在每次 AIMPs 输入后计算。

(6.84) 式中的第三项，核实投影 (位移) 算符 $\hat{P}(i)$ 保留原有形式，即 (6.83) 式不变。算符中的核实轨道 $|\varphi_{\mathrm{c}}\rangle$ 用一组足够大的全电子基组来表达。对一分量的相对论模型势，AIMPs 方法选择的参考态能量来自基于 Cowan-Griffin (CG) Hamilton 算符的全电子计算。为考虑旋–轨耦合效应，AIMPs 方法引入了全电子的 Wood-Boring (WB) 旋–轨耦合算符[69, 2]

$$\Delta\hat{V}_{\mathrm{cv,so}}^{\lambda}(\boldsymbol{r}_i)=\sum_l\left(\sum_k\frac{B_{lk}^{\lambda}}{r_{\lambda i}^2}\mathrm{e}^{-\beta_{lk}^{\lambda}r_{\lambda i}^2}\right)\hat{P}_l^{\lambda}(i)\hat{\boldsymbol{L}}_{\lambda i}\cdot\hat{\boldsymbol{S}}_i\hat{P}_l^{\lambda}(i) \tag{6.89}$$

式中，$\hat{\boldsymbol{l}}_{\lambda i}$, $\hat{\boldsymbol{S}}_i$, $\hat{P}_l^{\lambda}$ 的含义和赝势旋–轨耦合算符 (6.49) 式中对应各项相同。系数 B_{lk}^{λ} 和指数 β_{lk}^{λ} 则通过与全电子 WB 旋–轨耦合算符的径向部分 (见第 5 章) 比较，以最小二乘法拟合得到。

所有 AIMPs 参数和配套基组读者可以从 Seijo 的个人网站 (http://www.uam.es/personal_pdi/ciencias/lseijo/) 下载。

6.12 有效核实势与严格的相对论 Hamilton 算符

目前为止，我们讲到的 ECPs 方法中的价电子模型 Hamilton 算符都是非相对论形式的，我们当然也可以采用相对论的价电子 Hamilton 算符，这样一来，计算量会显著增大，那么是否也会显著提高计算结果的精确度呢？

1. 赝势和严格的相对论 Hamilton 算符

在 20 世纪 80~90 年代，人们陆续作了一些这方面的研究。例如，Datta 等[97] 使用 PK 赝势对 PbO 进行了四分量计算；Ishikawa 和 Malli[98] 研究了半局域赝势在四分量原子有限差分 DHF 计算中的应用；Dolg[99] 采用能量优化的 PPs 和 CPPs 对碱金属及其氢化物进行了四分量的 DHF 和 CI 计算；Pyper[100] 通过对四分量赝势 Hamilton 算符实施 FW 变换推导出了复杂的赝势表达式并将其用于原子计算[101]；等等。他们的研究表明，伪价电子轨道对直接相对论效应的贡献很小，对计算精度的提高极为有限，再考虑到高昂的相对论计算成本，赝势方法还是更适合与非相对论的价电子模型 Hamilton 算符同时使用。

2. 模型势和严格的相对论 Hamilton 算符

从第 1 章我们已经知道，粒子速度越大，相对论效应越明显。所以近核高速运动的内层电子的相对论效应要高于外层电子，是直接相对论效应的主要来源。与赝势相比，模型势保留了全电子价电子轨道的节点结构，因而更适合和严格的相对论 Hamilton 算符相结合。在过去的二十年间，有不少研究人员进行了这方面的尝试，报道的结果总的来说也不错，但实际应用并不多。

1995 年 Wittborn 和 Wahlgren[102] 发表了第三周期过渡金属的相对论 AIMPs，价层空间包括 4s, 4p, 4d, 5s, 5p, 5d, 6s 等轨道。他们将核实对价电子轨道的间接相对论效应包含在 MPs 中，直接的相对论效应则用自旋平均的相对论二级 DKH (DKH2) Hamilton 算符来严格处理。对原子 $s^2d^n \rightarrow s^1d^{n+1}$ 基态跃迁能量的 HF 计算表明，与 AE DKH2 的结果相比，AIMPs 计算结果的均方差为 0.05 eV。他们还计算了过渡金属一氢化物的分子结构，发现 AIMPs 得到的 M— H (M=La, Hf, Ir, Os, Pt) 的键长比 AE DKH2 计算值偏大，最多可达 0.04 Å。

对第一、二、三周期的过渡金属，Rakowitz 等[103, 104] 发表了与 DKH2 配套使用的自旋平均 AIMPs，他们选择了两种价层空间，一种较小，仅包括最外层 s 轨道和次外层的 p, d 轨道，即 $(n-1)$p $(n-1)$d ns；另一种较大，包括最外层 s 轨道和次外层的 s, p, d 轨道，即 $(n-1)$s $(n-1)$p $(n-1)$d ns。通过对第 5, 10 族过渡

金属氧化物 (VO, NbO, TaO, NiO, PdO, PtO) 的计算，他们发现后者 (价层空间较大) 得到的结果和 AE DKH 的结果更一致。

对镧系[105] 和锕系元素[106]，Seijo, Hirao 等开发了与 DKH3 配套使用的自旋平均 AIMPs，相应的旋–轨耦合算符则采用了与 AE DKH 相同的处理方法 (见第 5 章) —— 基于组态相互作用的原子平均场近似方法[107]。对镧系元素，选择的价层轨道包括 4f, 5s, 5p, 5d 和 6s；对锕系元素，有大小两种核实选择，大核实是将 5f, 6s, 6p, 6d, 7s 5 个轨道选为价层轨道；小核实的价层空间包括 5d, 5f, 6s, 6p, 6d, 7s 6 个价层轨道。通过比较 AE 和 MP 轨道能量、轨道半径期望值以及对 CeO, ThO 分子的计算结果，他们认为不需要将锕系元素中的 5d 轨道作价电子处理，即镧系和锕系元素可采用拥有相同价层电子数目的核实模型。必须指出的是在他们的测试中没有考虑到 f 轨道占据电子数发生变化的情形。

6.13　基于密度泛函理论的有效核实势

到目前为止，我们都是采用基于波函数理论的全电子 HF (DHF) 方法来获得优化赝势 (或模型势) 所需的参考态，由此得到的有效核实势又称从头算有效核实势。尽管如此，从头算赝势(尤其是小核实赝势) 用于密度泛函计算也获得了很大的成功。我们当然也可以直接从密度泛函理论 (DFT) 出发，得到基于 DFT 的有效核实势，其价电子基组通常采用平面波函数形式，广泛用于固体计算，是有效核实势方法中的一个重要分支[108−112]。用 DFT 或用从头算结果优化有效核实势的原理和步骤是一样的，区别在于二者对相关方程的表达形式不同，如在讨论 DFT 模型势时，就必须从 Kohn-Sham (KS) 方程出发重写 Huzinaga-Cantu 方程、对 DFT 赝势，伪价电子轨道要改用 KS 轨道等。下面我们简单介绍基于 DFT 的模型势和形状一致的赝势。

6.13.1　模型势

从头算有效核实势大多选择 HF (DHF) 计算全电子参考态，而密度泛函的种类很多，可选择的范围很广。下面我们介绍 Andzelm, Radzio, Salahub 发表的基于 LSD-VWN 泛函[113] 优化的模型势[108]。

假定一个自旋极化的体系，有 n_v 个价电子，且含自旋 σ $(\sigma = +,-)$ 的价电子轨道 $\varphi_\mathrm{v}^\sigma$ 和核实电子轨道 $\varphi_\mathrm{c}^\sigma$ 正交。从 Kohn-Sham (KS) 方程出发，将 Huzinaga-Cantu 方程 (6.79) 式重写为

$$\left[\hat{F}^\sigma + \sum_\mathrm{c}(-2\epsilon_\mathrm{c}^\sigma)|\varphi_\mathrm{c}^\sigma\rangle\langle\varphi_\mathrm{c}^\sigma|\right]|\varphi_\mathrm{v}^\sigma\rangle = \epsilon_\mathrm{v}^\sigma|\varphi_\mathrm{v}^\sigma\rangle \tag{6.90}$$

式中

$$\hat{F}^{\sigma} = \hat{F}_{\mathrm{v}}^{\sigma} + \Delta\hat{V}^{\sigma,\mathrm{MP}} \tag{6.91}$$

$$\hat{F}_{\mathrm{v}}^{\sigma}(\boldsymbol{r}) = -\frac{1}{2}\nabla^2 - \sum_{\lambda}\frac{Q_{\lambda}}{r_{\lambda}} + \int\frac{\rho_{\mathrm{v}}(\boldsymbol{r}')d\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|} + v_{\mathrm{xc}}\left[\rho_{+,v},\rho_{-,v}\right] \tag{6.92}$$

式中，Q_{λ} 是核实 λ 所带电荷；$\rho_{+,v}$ 和 $\rho_{-,v}$ 分别代表价轨道自旋为上和自旋为下的密度，$\rho_{\mathrm{v}} = \rho_{+,v} + \rho_{-,v}$。$v_{\mathrm{xc}}$ 表示交换相关势。密度 $\rho_{\sigma,v}$ 定义为

$$\rho_{\sigma,v}(\boldsymbol{r}) = \sum_{i}^{n_{\mathrm{v}}} f_i^{\sigma}|\varphi_i^{\sigma}(\boldsymbol{r})|^2 \tag{6.93}$$

式中，f_i^{σ} 是布居数。定义模型势 $\hat{V}^{\sigma,\mathrm{MP}}$

$$\hat{V}^{\sigma,\mathrm{MP}} = \Delta\hat{V}^{\sigma,\mathrm{MP}} + \hat{P} \tag{6.94}$$

$$\hat{P} = \sum_{\mathrm{c}}(-2\epsilon_{\mathrm{c}}^{\sigma})|\varphi_{\mathrm{c}}^{\sigma}\rangle\langle\varphi_{\mathrm{c}}^{\sigma}| \tag{6.95}$$

利用核实冻结近似，假定各原子的核实轨道并不重叠，忽略交换势和相关势的交叉项等，于是 $\hat{V}^{\sigma,\mathrm{MP}}$ 可以写成所有原子的模型势 $\hat{V}_{\lambda}^{\sigma,\mathrm{MP}}$ 之和

$$\hat{V}^{\sigma,\mathrm{MP}} = \sum_{\lambda}\hat{V}_{\lambda}^{\sigma,\mathrm{MP}} = \sum_{\lambda}\left(\Delta\hat{V}_{\lambda}^{\sigma,\mathrm{MP}} + \hat{P}_{\lambda}\right) \tag{6.96}$$

式中

$$\Delta\hat{V}_{\lambda}^{\sigma,\mathrm{MP}} = -\frac{n_{\mathrm{c}}^{\lambda}}{r_{\lambda}} + \int\frac{\rho_{\mathrm{c}}^{\lambda}(\boldsymbol{r}')d\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|} + v_{\mathrm{xc}}\left[\rho_{+,\mathrm{c}}^{\lambda},\rho_{-,\mathrm{c}}^{\lambda}\right] \tag{6.97}$$

$\Delta\hat{V}_{\lambda}^{\sigma,\mathrm{MP}}$ 的分析表达式可以写为 [108]

$$\Delta\hat{V}_{\lambda}^{\sigma,\mathrm{MP}} = \sum_{k}A_k^{\lambda}\frac{\mathrm{e}^{-\alpha_k^{\lambda}r^2}}{r} \tag{6.98}$$

与从头算 MPs 对展开系数的要求一样 (见 6.11.3 节)，上式中系数 A_k^{λ} 满足 $\sum\limits_{k}A_k^{\lambda} = Z^{\lambda} - Q^{\lambda} = n_{\mathrm{c}}^{\lambda}$。投影算符 $\hat{P}_{\lambda}$ 中的核实轨道 (见 (6.94) 式) 用一组 Gauss 函数展开以最小二乘法拟合得到。全电子原子轨道则来自于基于 CG/WB Hamilton 算符[21]的 LSD-VWN 原子有限差分计算。

6.13.2 形状一致的赝势

在 6.9.2 节我们学习了如何得到从头算形状一致赝势。采用类似步骤我们同样能得到基于 DFT 理论优化的形状一致赝势。二者的区别在于，对给定的原子态，与真实的全电子轨道能量 ($\epsilon_{\mathrm{v},lj}$) 对应的 DFT 伪价电子轨道 ($\varphi_{\mathrm{p},lj}$) 是 KS 轨道。为便于所得赝势用于固体计算，必须考虑束缚态和散射分波 (即非束缚态) 两种情况。为此，Hamann, Schlüter 和 Chiang (HSC) 提出伪分波 $\varphi_{\mathrm{p},lj}$ 必须具有模守恒 (norm-conserving) 性质[114]：

(1) 全电子和伪价电子波函数的本征值相同。

(2) 当 $r > r_c$ 时全电子价电子波函数和赝势的伪价电子波函数相同。

(3) 当 $r > r_c$ 时全电子和赝势价电子电荷积分 (从 0 到 r) 值相同。

(4) 当 $r > r_c$ 时全电子和赝势的价电子分波的对数导数以及它们的能量一阶导数相同。

对周期表中从 H 到 Pu 的所有元素, Bachelet, Hamann, Schlüter[110] 发表了基于局域密度泛函的相对论赝势。如果只考虑束缚态，模守恒条件可以简化为从头算形状一致赝势的相应条件 (见 (6.66) 式)。在只考虑束缚态的条件下，对周期表中从 H 到 Am 的所有元素 Delley 发表了基于局域轨道方法的密度泛函半核实赝势 (density functional semi-core，DSPP)[111]。DSPP 选择了 2~3 个相关原子的离子态作为参考态，用于 AE 参考态计算的密度泛函主要为 PBE (Perdew-Burke-Ernzerhof) 梯度校正泛函以及 LDA (local density approximation) 条件下的 PW (Perdew-Wang) 局域相关等。为了提高计算的稳定性以及校正核实冻结近似， Delley 等引入了核实密度模型 (又称非线性核实校正)[111, 115]。 Delley[111] 发现由赝势近似导致的误差要小于由密度泛函引入的误差。

参 考 文 献

[1] Klobukowski M, Huzinaga S, Sakai Y. Model core potentials: Theory and application. // Leszczynski J. Computional Chemistry: Reviews of Current Trends. Vol 3. Singapore: World Scientific, 1999: 49~74.

[2] Seijo L, Barandiarán Z. The *ab initio* model potential method: a common strategy for effective core potential and embedded cluster calculations.// Leszczynski J. Computional Chemistry: Reviews of Current Trends. Vol 4. Singapore: World Scientific, 1999: 55~152.

[3] Pyykkö P, Stoll H. Relativistic pseudopotential calculations, 1993~1999. // Hincliffe A. RSC Spec Period Rep. Chemical Modelling, Applications and Theory. Vol 1. Cambridge: Royal Soc of Chemistry, 2000: 239~305.

[4] Schwerdtfeger P. Relativistic pseudopotentials. // Kaldor U, Wilson S. Progress in Theoretical Chemistry and Physics - Theoretical Chemistry and Physics of Heavy and Superheavy Element. Dordrecht: Kluwer Academic publisher, 2003: 399~438.

[5] Cao X, Dolg M. Relativistic pseudopotentials.// Barysz M , Ishikawa Y. Relativistic Methods for Chemists. Challenges and Advances in Computational Physics. Vol 10. Dordrecht: Springer Science+Business Media B V , 2010: 215~277.

[6] Dolg M, Cao X. Relativistic pseudopotentials: Their development and scope of applications. Chem Rev, 2012, 112: 403~480.

[7] Hellmann H. A new approximation method in the problem of many electron electrons. J Chem Phys, 1935, 3: 61.

[8] Phillips J C, Kleinman L. New method for calculating wave functions in crystals and molecules. Phys Rev, 1959, 116: 287～294.

[9] Smallwood C J, Larsen R E, Glover W J, et al. A computationally efficient exact pseudopotential method I. Analytic reformulation of the Phillips-Kleinman theory. J Chem Phys, 2006, 125: 074102-1～074102-9.

[10] Smallwood C J, Meija C N, Glover W J, et al. A computationally efficient exact pseudopotential method II. Application to the molecular pseudopotential of an excess electron interacting with tetrahydrofuran (THF). J Chem Phys, 2006, 125: 074103-1～074103-9.

[11] Weeks J D, Rice S A. Use of pseudopotentials in atomic-structure calculations. J Chem Phys, 1968, 49: 2741～2755.

[12] Barthelat J C, Durand P. Recent progress of pseudo-potential methods in quantum chemistry. Gaz Chim Ital, 1978, 108: 225～236.

[13] Abarenkov I V, Heine V. The model potential for positive ions. Phil Mag, 1965, 12: 529～537.

[14] Schwarz W H E. Das kombinierte Näherungsverfahren. Theor Chim Acta, 1968, 11: 307～324.

[15] Kahn L R, Goddard W A. A direct test of the validity of the use of pseudopotentials in molecules. Chem Phys Lett, 1968, 2: 667～670.

[16] Kahn L R, Goddard W A. *Ab initio* effective potentials for use in molecular calculations. J Chem Phys, 1972, 56: 2685～2701.

[17] Kahn L R, Baybutt P, Truhlar D G. *Ab initio* effective core potentials: Reduction of all-electron molecular structure calculations to calculations involving only valence electrons. J Chem Phys, 1976, 65: 3826～3853.

[18] Durand P, Barthelat J C. A theoretical method to determine atomic pseudopotentials for electronic structure calculations of molecules and solids. Theor Chim Acta, 1975, 38: 283～302.

[19] Christiansen P A, Lee Y S, Pitzer K S. Improved *ab initio* effective core potentials for molecular calculations. J Chem Phys, 1979, 71: 4445～4450.

[20] Kahn L R, Hay P J, Cowan R D. Relativistic effects in *ab initio* effective core potentials for molecular calculations — applications to uranium atom. J Chem Phys, 1978, 68: 2386～2397.

[21] Cowan R D, Griffin D C. Approximate relativistic corrections to atomic radial wave functions. J Opt Soc Am, 1976, 66: 1010～1014.

[22] Hafner P, Schwarz W H E. Pseudopotential approach including relativistic effects. J Phys B: At Mol Phys, 1978, 11: 217～233.

[23] Lee Y S, Ermler W C, Pitzer K S. *Ab initio* effective core potentials including relativistic effects I formalism and applications to the Xe and Au Atoms. J Chem Phys, 1977, 67: 5861~5876.

[24] Ermler W C, Lee Y S, Christiansen P A, et al. *Ab initio* effective core potentials including relativistic effects A procedure for the inclusion of spin-orbit coupling in molecular wavefunctions. Chem Phys Lett, 1981, 81: 70~74.

[25] Pitzer R M, Winter N W. Electronic-structure methods for heavy-atom molecules. J Phys Chem, 1988, 92: 3061~3063.

[26] Christiansen P A, Balasubramanian K, Pitzer K S. Relativistic *ab initio* molecular structure calculations including configuration interaction with application to six states of TlH. J Chem Phys, 1982, 76: 5087~5092.

[27] DiLabio G A, Christiansen P A. Low-lying 0^+ states of bismuth hydride. Chem Phys Lett, 1997, 277: 473~477.

[28] Esser M. Direct MRCI method for the calculation of relativstic many-electron wavefunctions 1 general formalism. Int J Quant Chem, 1984, 16: 313~338.

[29] Sjovoll M, Gropen O, Olsen J. A determinantal approach to spin-orbit configuration interaction. Theor Chem Acc, 1997, 97: 301~312.

[30] Vallet V, Maron L, Teichteil C, et al. A two-step uncontracted determinantal effective Hamiltonian-based SO-CI method. J Chem Phys, 2000, 113: 1391~1402.

[31] Hess B A, Marian C M, Wahlgren U, et al. A mean-field spin-orbit method applicable to correlated wavefunctions. Chem Phys Lett, 1996, 251: 365~371.

[32] Fuentealba P, Preuss H, Stoll H, et al. A proper account of core-polarization with pseudopotentials: single valence-electron alkali compounds. Chem Phys Lett, 1982, 89: 418~422.

[33] Müller W, Flesch J, Meyer W. Treatment of intershell correlation effects in *ab initio* calculations by use of core polarization potentials method and application to alkali and alkaline earth atoms. J Chem Phys, 1984, 80: 3297~3310.

[34] (a) Schwerdtfeger P, Silberbach H. Multicenter integrals over long-range operators using Cartesian Gaussian functions. Phys Rev A, 1988, 37: 2834~2842.
(b) Schwerdtfeger P, Silberbach H. Erratum: ibidem. Phys Rev A, 1990, 42: 665.

[35] Smit M J. Multicenter integrals over polarization potential operators. Int J Quant Chem, 1999, 73: 403~416.

[36] Igel-Mann G, Stoll H, Preuss H. Pseudopotentials for main group elements (IIIa through VIIa). Mol Phys, 1988, 65: 1321~1328.

[37] Igel-Mann G, Stoll H, Preuss H. Pseudopotential study of monohydrides and monoxides of main group elements K through Br. Mol Phys, 1988, 65: 1329~1336.

[38] Weigand A, Cao X, Dolg M. Quasirelativistic f-in-core pseudopotentials and core-polarization potentials for trivalent actinides and lanthanides: Molecular test for tri-

fluorides. Theor Chem Acc, 2010, 126: 117~127.

[39] Stoll H, Fuentealba P, Dolg M, et al. Cu and Ag as one-valence-electron atoms: Pseudopotential results for Cu_2, Ag_2, CuH, AgH, and the corresponding cations. J Chem Phys, 1983, 79: 5532~5542.

[40] Igel G, Wedig U, Dolg M, et al. Cu and Ag as one-valence-electron atoms: Pseudopotential CI results for CuO and AgO. J Chem Phys, 1984, 81: 2737~2740.

[41] Stoll H, Fuentealba P, Schwerdtfeger P, et al. Cu and Ag as one-valence-electron atoms: CI results and quadrupole corrections for Cu_2, Ag_2, CuH, AgH. J Chem Phys, 1984, 81: 2732~2736.

[42] Preuss H. Untersuchungen zum kombinierten Näherungsverfahren. Z Naturf, 1955, 10a: 365~373.

[43] Preuss H, Stoll H, Wedig U, et al. A combination of pseudopotentials and density functionals. Int J Quant Chem, 1981, 19: 113~130.

[44] Cao X, Dolg M. Relativistic energy-consistent *ab initio* pseudopotentials as tools for quantum chemical investigations of actinide systems. Coord Chem Rev, 2006, 250: 900~912.

[45] Dolg M, Cao X. Accurate relativistic small-core pseudopotentials for actinides energy adjustment for uranium and first applications to uranium hydride. J Phys Chem A, 2009, 113: 12573~12581.

[46] Moritz A, Cao X, Dolg M. Quasirelativistic energy-consistent 5f-in-core pseudopotentials for divalent and tetravalent actinide elements. Theor Chem Acc, 2007, 118: 2763~2774.

[47] Moritz A, Cao X, Dolg M. Quasirelativistic energy-consistent 5f-in-core pseudopotentials for trivalent actinide elements. Theor Chem Acc, 2007, 117: 473~481.

[48] Dyall K G, Grant I P, Johnson C T, et al. GRASP: a general-purpose relativistic atomic structure program. Comput Phys Commun, 1989, 55: 425~456.

[49] Wood J H, Boring A M. Improved Pauli Hamiltonian for local-potential problems. Phys Rev B, 1978, 18: 2701~2711.

[50] Küchle W, Dolg M, Stoll H, et al. Energy-adjusted pseudopotentials for the actinides parameter sets and test calculations for thorium and thorium monoxide. J Chem Phys, 1994, 100: 7535~7542.

[51] Froese-Fischer C. The Hartree-Fock Method for Atoms. New York: John Wiley & Sons Inc, 1977.

[52] Hess B A, Dolg M. Relativistic quantum chemistry with pseudopotentials and transformed Hamiltonians. // Hess B A. Relativistic Effects in Heavy-Element Chemistry and Physics. Wiley Series in Theoretical Chemistry. Vol 12. New York: John Wiley & Sons Inc, 2002: 89~122.

[53] Blaudeau J P, Curtiss L A. Optimized Gaussian basis sets for use with relativistic

effective (core) potentials: K, Ca, Ga-Kr. Int J Quant Chem, 1997, 61: 943~952.

[54] Ermler W C, Ross R B, Christiansen P A. Spin-orbit coupling and other relativistic effects in atoms and molecules. Adv Quantum Chem, 1988, 19: 139~182.

[55] Ermler W C, Ross R B, Christiansen P A. *Ab initio* relativistic effective potentials with spin-orbit operators. Ⅵ. Fr through Pu. Int J Quant Chem, 1991, 40: 829~846.

[56] Hurley M M, Pacios L F, Christiansen P A, et al. *Ab initio* relativistic effective potentials with spin-orbit operators. Ⅱ. K through Kr. J Chem Phys, 1986, 84: 6840~6853.

[57] LaJohn L A, Christiansen P A, Ross R B, et al. *Ab initio* relativistic effective potentials with spin-orbit operators. Ⅲ. Rb through Xe. J Chem Phys, 1987, 87: 2812~2824.

[58] Pacios L F, Christiansen P A. *Ab initio* relativistic effective potentials with spin-orbit operators. Ⅰ. Li through Ar. J Chem Phys, 1985, 82: 2664~2671.

[59] Nash C S, Bursten B E, Ermler W C. *Ab initio* relativistic potentials with spin-orbit operators. Ⅶ. Am through element 118. J Chem Phys, 1994, 106: 5133~5142.

[60] Ross R B, Powers J M, Atashroo T, et al. *Ab initio* relativistic effective potentials with spin-orbit operators. Ⅳ. Cs through Rn. J Chem Phys, 1990, 93: 6654~6670.

[61] Ross R B, Gayen S, Ermler W C. *Ab initio* relativistic effective potentials with spin-orbit operators. Ⅴ. Ce through Lu. J Chem Phys, 1994, 100: 8145~8155.

[62] Wildman S A, DiLabio G A, Christiansen P A. Accurate relativistic effective potentials for the sixth-row main group elements. J Chem Phys, 1997, 107: 9975~9979.

[63] Hay P J, Wadt W R. *Ab initio* effective core potentials for molecular calculations potentials for the transition metal atoms Sc to Hg. J Chem Phys, 1985, 82: 270~282.

[64] Hay P J, Wadt W R. *Ab initio* effective core potentials for molecular calculations potentials for K to Au including the outermost core orbitals. J Chem Phys, 1985, 82: 299~310.

[65] Wadt W R, Hay P J. *Ab initio* effective core potentials for molecular calculations. potentials for main group elements Na to Bi. J Chem Phys, 1985, 82: 284~298.

[66] Roy L E, Hay P J, Martin R L. Revised basis sets for the LANL effective core potentials. J Chem Theor Comput, 2008, 4: 1029~1031.

[67] Barthelat J C, Durand P, Serafini A. Non-empirical pseudopotentials for molecular calculations. I. The PSIBMOL algorithm and test calculations. Mol Phys, 1977, 33: 159~180.

[68] Cundari T R, Stevens W J. Effective core potential methods for the lanthanides. J Chem Phys, 1993, 98: 5555~5565.

[69] Stevens W J, Basch H, Krauss M. Compact effective potentials and efficient shared-exponent basis sets for the first- and second-row atoms. J Chem Phys, 1984, 81: 6026~6033.

[70] Stevens W J, Krauss M, Basch H, et al. Relativistic compact effective potentials and efficient, shared-exponent basis sets for the third-, fourth-, and fifth-row atoms. Can

J Chem, 1992, 70: 612∼630.

[71] Titov A V, Mosyagin N S. Generalized relativistic effective core potential: Theoretical ground. Int J Quant Chem, 1999, 71: 359∼401.

[72] Mosyagin N S, Titov A V, Latajka Z. Generalized relativistic effective core potential: Gaussian expansion of potentials and pseudospinors for atoms Hg through Rn. Int J Quant Chem, 1997, 63: 1107∼1122.

[73] Titov A V, Mosyagin N S. The generalized relativistic effective core potential method: Theory and calculations. Russ J Phys Chem, 2000, 74: S376∼S387.

[74] Titov A V, Mosyagin N S, Isaev T A, et al. Accuracy and efficiency of modern methods for electronic structure calculation on heavy- and superheavy-elements compounds. Phys At Nucl, 2003, 66: 1188∼1198.

[75] Mosyagin N S, Petrov A N, Titov A V, et al. Generalized RECP accounting for Breit effects: Uranium, plutonium and superheavy elements 112, 113, 114. // Julien J P, Maruani J, Mayou D. Recent Advances in the Theory of Chemical and Physical Systems. Progress in Theoretical Chemistry and Physics, Part II. Vol 15, Berlin: Springer-Verlag, 2006: 253∼284.

[76] Petrov A N, Mosyagin N S, Titov A V, et al. Accounting for Breit interaction in relativistic effective core potential calculations of actinides. J Phys B At Mol Opt Phys, 2004, 23: 4621∼4637.

[77] Zaitsevskii A V, Wüllen C V, Titov A V. Relativistic pseudopotential model for superheavy elements: applications to chemistry of eka-Hg and eka-Pb. Rus Chem Rev, 2009, 12: 1173∼1181.

[78] Zaitsevskii A V, Wüllen C V, Rykova E A, et al. Two-component relativistic density functional theory modeling of the adsorption of element 114 (eka-lead) on gold. Phys Chem Chem Phys, 2010, 16: 4152∼4156.

[79] Titov A V, Mosyagin N S, Petrov A N, et al. Two-step method for precise calculation of core properties in molecules. Int J Quant Chem, 2005, 104: 223∼239.

[80] Evarestov R A, Losev M V, Panin A I, et al. Electronic structure of crystalline uranium nitride: LCAO DFT calculations. Phys Stat Sol B, 2008, 245: 114∼122.

[81] Petrov A N, Mosyagin N S, Titov A V, et al. *Ab initio* study of Hg-Hg and E112-E112 van der Waals interactions. Phys At Nucl, 2009, 72: 396∼400.

[82] Almlöf J, Taylor P R. General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. J Chem Phys, 1987, 86: 4070∼4077.

[83] Almlöf J, Helgaker T, Taylor P R. Gaussian basis sets for high-quality *ab initio* calculations. J Phys Chem, 1988, 92: 3029∼3033.

[84] Dunning T H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J Chem Phys, 1989, 90: 1007∼1023.

[85] Weigend F, Ahlrichs R. Balanced basis sets of split valence, triple zeta valence and

quadruple zeta valence quality for H to Rn:Design and assessment of accuracy. Phys Chem Chem Phys, 2005, 7: 3297~3305.

[86] Weigend F, Baldes A. Segmented contracted basis sets for one- and two-component Dirac-Fock effective core potentials. Phys Chem Chem Phys, 2010, 133: 174102-1~174102-10.

[87] Martin J M, Sundermann A. Correlation consistent valence basis sets for use with the Stuttgart-Dresden-Bonn relativistic effective core potentials: The atoms Ga–Kr and In–Xe. J Chem Phys, 2001, 114: 3408~3420.

[88] Peterson K A, Figgen D, Goll E, et al. Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements. J Chem Phys, 2003, 119: 11113~11123.

[89] Peterson K A, Yousaf K E. Molecular core-valence correlation effects involving the post-d elements Ga-Rn: Benchmarks and new pseudopotential-based correlation consistent basis sets. J Chem Phys, 2010, 133: 174116-1~174116-8.

[90] Bonifacic V, Huzinaga S. Atomic and molecular calculations with the model potential method I. J Chem Phys, 1974, 60: 2779~2786.

[91] Huzinaga S, Cantu A A. Theory of separability of many-electron systems. J Chem Phys, 1971, 55: 5543~554.

[92] Huzinaga S, McWilliams D, Cantu A A. Projection operators in Hartree-Fock theory. Adv Quantum Chem, 1972, 7: 187~220.

[93] Huzinaga S, Seijo L, Barandiarán Z, et al. The *ab initio* model potential method. Main group elements. J Chem Phys, 1987, 86: 2132~2145.

[94] Katsuki S, Huzinaga S. An effective Hamiltonian method for valence-electron molecular calculations. Chem Phys Lett, 1988, 152: 203~206.

[95] Sakai Y. New developments in the model potential method. J Chem Phys, 1981, 75: 1303~1308.

[96] Sakai Y, Huzinaga S. The use of model potentials in molecular calculations I. J Chem Phys, 1982, 86: 2132~2145.

[97] Datta S N, Ewig C S, van Wazer J R. Application of effective potentials to relativistic Hartree-Fock calculations. Chem Phys Lett, 1978, 57: 83~89.

[98] Ishikawa Y, Malli G. Effective core potentials for fully relativistic Dirac-Fock calculations. J Chem Phys, 1981, 75: 5423~5431.

[99] Dolg M. Fully relativistic pseudopotentials for alkaline atoms Dirac-Hartree-Fock and configuration interaction calculations of alkaline monohydrides. Theor Chim Acta, 1996, 93: 141~156.

[100] (a) Pyper N C. Relativistic pseudopotential theories and corrections to the Hartree-Fock method. Mol Phys, 1980, 39: 1327~1358.
(b) Pyper N C. Correction. Mol Phys, 1980, 41: 949~951.

[101] Pyper N C, Marketos P. 1st order perturbation treatment for relativistic pseudopotential and corrections to the Hartree-Fock method 2 results. Mol Phys, 1981, 42: 1073～1103.

[102] Wittborn C, Wahlgren U. New relativistic effective core potentials for heavy elements. Chem Phys, 1995, 201: 357～362.

[103] Rakowitz F, Marian C M, Seijo L, et al. Spin-free relativistic no-pair *ab initio* core model potentials and valence basis sets for the transition metal elements Sc to Hg I. J Chem Phys, 1999, 110: 3678～3686.

[104] Rakowitz F, Marian C M, Seijo L. Spin-free relativistic no-pair *ab initio* core model potentials and valence basis sets for the transition metal elements Sc to Hg II. J Chem Phys, 1999, 111: 10436～10443.

[105] Tsuchiya T, Nakajima T, Hirao K, et al. A third-order Douglas-Kroll *ab initio* model potential for the lanthanides. Chem Phys Lett, 2002, 361: 334～340.

[106] Paulovič J, Nakajima T, Hirao K, et al. Third-order Douglas-Kroll *ab initio* model potential for actinide elements. J Chem Phys, 2002, 117: 3597～3604.

[107] Paulovič J, Nakajima T, Hirao K, et al. Relativistic and correlated calculation on the ground and excited states of ThO. J Chem Phys, 2003, 119: 798～805.

[108] Andzelm J, Radzio E, Salahub D R. Model potential calculations for second-row transition metal molecules within the local-spin-density method. J Chem Phys, 1985, 83: 4573～4580.

[109] Bachelet G B, Schlüter M. Relativistic norm-conserving pseudopotentials. Phys Rev B, 1982, 25: 2103～2108.

[110] Bachelet G B, Hamann D R, Schlüter M. Pseudopotentials that work: From H to Pu. Phys Rev B, 1982, 26: 4199～4228.

[111] Delley B. Hardness conserving semilocal pseudopotentials. Phys Rev B, 2002, 66: 155125-1～155125-9.

[112] Gutsev G L. Numerical pseudopotentials within DV-X_α framework. Adv Quant Chem, 1997, 29: 137～157.

[113] Vosko S H, Wilk L, Nusair M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. Can J Phys, 1980, 58: 1200～1211.

[114] Hamann D R, Schlüter M, Chiang C. Norm-conserving pseudopotentials. Phys Rev Lett, 1979, 43: 1494～1497.

[115] Louie S G, Froyen S, Cohen M L. Nonlinear ionic pseudopotentials in spin-density-functional calculations. Phys Rev B, 1982, 26: 1738～1742.

第 7 章　元素周期表中的相对论效应

1. 直接相对论效应

以类氢原子 Yb^{69+}(核电荷数 Z 为 70) 为例，其 1s 电子的非相对论速度约为 70 a.u.①，大约是光速 (137 a.u.) 的 51%。由第 1 章 (1.33) 式得 Yb^{69+} 中 1s 电子的相对论质量 (m_{1s}) 与静止质量 (m_0) 的关系为

$$m_{1s} = \frac{m_0}{\sqrt{1-(Z/c)^2}} = 1.16\ m_0 \tag{7.1}$$

由 Bohr 半径 (a_0) 表达式

$$a_0 = \frac{\hbar^2}{m_{1s}e^2} = \frac{0.86\hbar^2}{m_{1s}e^2} \tag{7.2}$$

知，与非相对论的 Bohr 轨道半径相比，考虑相对论效应后 Yb^{69+} 的 1s 轨道半径收缩了大约 14%。再来看电子的能量，Yb^{69+} 的 1s 电子的非相对论能量可写为

$$E_{1s} = -\frac{1}{2}Z^2 = -2450\ \text{Hartree} \tag{7.3}$$

式中，Hartree 为原子单位制中的能量单位。由第 2 章 (2.140) 式可得 1s 的相对论能量 (略去静止能量 m_0c^2)

$$E_{1s} = -\frac{1}{2}Z^2 - \frac{Z^4}{8c^2} = -2610\ \text{Hartree} \tag{7.4}$$

与 (7.3) 式相比，1s 电子的相对论能量明显低于非相对论能量。一般地，将考虑相对论效应后轨道能量降低、轨道半径减少的现象称为“相对论收缩”。

对更外层的 ns 轨道，由轨道正交性条件

$$\int_0^\infty R_{1s}(r)R_{ns}(r)r^2\mathrm{d}r = 0 \tag{7.5}$$

知其他 ns 轨道也必然向内收缩。不仅 s 轨道，而且 p 轨道以及最内层的 d 轨道也会发生相对论收缩。一般将相对论收缩和自旋–轨道分裂并称为“直接相对论效应”。

2. 间接相对论效应

由于 s, p 轨道的相对论收缩，内层轨道能更好地屏蔽核电荷，降低了有效核电荷对外层电子的吸引力，从而导致了外层 d, f 轨道的膨胀，也称相对论膨胀，即

① 1 a.u. = 2.187 7×10^6m·s^{-1}

轨道能量升高、轨道半径增大。例如，考虑相对论效应后，Yb 的基态 4f 轨道能量为非相对论能量的 70% 左右，轨道半径膨胀了约 5%[1]。由于膨胀是由内层轨道收缩引起的，因此也称为间接相对论效应。

由于不同原子中包含电子数目的差异，直接相对论效应和间接相对论效应对原子性质的影响和大小不尽相同，接下来我们就按照周期表中元素分类，分别介绍相对论效应对化学性质的影响。

7.1 主族元素

主族元素包括元素周期表中的第 1, 2, 13~18 族元素，这些元素的共同特点是它们的价层电子结构只包含角量子数较低的 s 和 (或) p 轨道。

1. 第 1, 2 族元素

这两族元素又称为碱金属和碱土金属族元素，它们的价电子基态组态分别为 $n\mathrm{s}^1$, $n\mathrm{s}^2$ (n = 2~7)，价层 s 轨道的相对论收缩是影响这些元素化学性质的主要相对论效应。如果仅考虑非相对论的原子壳结构膨胀效应，碱金属或碱土金属族元素的原子半径将随着原子序数的增加而逐渐增大，原子的第一电离能则逐渐降低。实验数据表明 (见表 7.1)，从第一周期到第五周期元素，第一电离能 (IP_1) 逐渐降低，从第六周期元素 (Cs，Ba) 到第七周期元素 (Fr，Ra)，IP_1 的值升高。这是由于随着原子序数的增加，相对论收缩效应逐渐增大，最后超过了原子的壳结构膨胀效应。

表 7.1 第 1, 2 族元素 ns 轨道的 Dirac-Fock 能量 ($-\epsilon_{\mathrm{rel}}$, eV)、相对论能量增加 ($\epsilon_{\mathrm{rel}}/\epsilon_{\mathrm{nrel}}$) 及第一电离能 ($\mathrm{IP}_1$, eV)

元素名称	n	$-\epsilon_{\mathrm{rel}}^{[1]}$	$\epsilon_{\mathrm{rel}}/\epsilon_{\mathrm{nrel}}^{[1]}$	$\mathrm{IP}_1^{[3]}$
Li	2	5.343	1.0001	5.392
Na	3	4.962	1.0013	5.139
K	4	4.028	1.0038	4.341
Rb	5	3.811	1.0185	4.177
Cs	6	3.490	1.0370	3.894
Fr	7	3.611	1.1255	4.073
Be	2	8.417	1.0002	9.323
Mg	3	6.897	1.0015	7.646
Ca	4	5.342	1.0040	6.113
Sr	5	4.932	1.0157	5.695
Ba	6	4.440	1.0359	5.212
Ra	7	4.523	1.1174	5.278

早在 1972 年，Fricke 和 Waber[2] 就发现，当原子序数大于 55 (Cs) 时，价层 s 轨道的相对论收缩效应就超过了原子的壳结构膨胀效应，Cs 因此成为自然界中最大

且最容易电离的原子[3]。但 Cs^+ 的离子半径并不是最大的，原因是当碱金属和碱土金属的外层 s 电子电离后，所得离子的离子半径主要由 $(n-1)$p 轨道决定，$(n-1)$p 轨道的相对论收缩要比 ns 轨道小得多，此时原子的壳结构膨胀效应重新回到主导地位，所以第七周期 Fr^+ 和 Ra^{2+} 的离子半径 (Fr^+：1.94 Å①；Ra^{2+}：1.64 Å) 要大于同族第六周期的 Cs^+ 和 Ba^{2+} 的离子半径 (Cs^+：1.81 Å；Ba^{2+}：1.49 Å)。离子的上述特点也会反映到分子结构上。例如，相对论计算表明，离子型化合物 XH_2 和 XH^+ (X=Ba, Ra) 中 Ra— H 键的键长要大于 Ba— H 键的键长[4]。

2. 第 13~18 族元素

第 13~18 族元素又分别称为硼族元素、碳族元素、氮族元素、氧族元素、卤素元素和惰性气体。它们的价电子组态为 $ns^2np^m(n=2\sim7, m=1\sim6)$。这里我们主要讨论 6p 轨道的自旋–轨道分裂，7p 轨道的旋–轨耦合效应将在 7.4 节详细讨论。

从表 7.2 可以看出，第六周期的 Pb, Bi, Po, At, Rn 等中性原子的 $6p_{1/2}$ 轨道能量比 $6p_{3/2}$ 至少低 1.50 eV，远大于同族较轻原子价层 p 轨道的自旋–轨道分裂。所以对第六周期原子，电子将优先填充 $6p_{1/2}$ 轨道，在 $6p_{1/2}$ 轨道充满 (两个电子) 后，电子再继续填充能量较高的 $6p_{3/2}$ 轨道。实验也测得第六周期原子的第一电离能只在 6p 轨道电子布居数从 1 增加到 2 时增加；而对其他周期的 13~18 族原子，当价层 p 轨道的电子布居数从 1 增加 2，再到 3 时，原子的第一电离能均升高。所以对第六周期原子，传统的 LS 耦合不再适用，在 jj 耦合图像下其基态价电子组态分别表示为：Tl，$6s^26p^1_{1/2}$；Pb，$6s^26p^2_{1/2}$；Bi，$6s^26p^2_{1/2}6p^1_{3/2}$；Po，$6s^26p^2_{1/2}6p^2_{3/2}$；At，$6s^26p^2_{1/2}6p^3_{3/2}$；

表 7.2 第 13~18 族元素基态的 $np(n=2\sim6)$ 轨道自旋–轨道分裂 (eV)[1]

	B	C	N	O	F	Ne
$2p_{3/2}$–$2p_{1/2}$	0.002 53	0.007 86	0.019 0	0.039 0	0.072 6	0.124
	Al	Si	P	S	Cl	Ar
$3p_{3/2}$–$3p_{1/2}$	0.014 6	0.031 2	0.056 4	0.092 3	0.141	0.206
	Ga	Ge	As	Se	Br	Kr
$4p_{3/2}$–$4p_{1/2}$	0.098 2	0.181	0.284	0.411	0.561	0.739
	In	Sn	Sb	Te	I	Xe
$5p_{3/2}$–$5p_{1/2}$	0.264	0.439	0.643	0.876	1.140	1.436
	Tl	Pb	Bi	Po	At	Rn
$6p_{3/2}$–$6p_{1/2}$	0.947	1.502	2.104	2.760	3.475	4.255

① 1 Å= 10^{-10} m

Rn，$6s^2 6p_{1/2}^2 6p_{3/2}^4$。由于 $6p_{1/2}$ 轨道能量过低，以上元素的主要化合价常与同族的其他元素不同。

通过对双原子分子 XH (X=B, Al, Ga, In, Tl) 的计算，Pyykkö 和 Desclaux 发现[5]，对 Tl— H 成键轨道的主要贡献来自于 $6p_{1/2}$，对其他硼族元素，X— H (X=B, Al, Ga, In) 成键轨道的主要贡献来自于价层 $p_{3/2}$ 轨道，所以 Tl 的主要化合价为 1，其他硼族元素的主要化合价为 3。由于 $6p_{1/2}$ 轨道的成键能力较弱，Tl— H 键的结合能 (2.18 eV) 要低于同族元素中的 In— H 键的结合能 (2.48 eV)[5]。

由表 7.2 可以看出，Pb 的 $6p_{1/2}$ 轨道比 $6p_{3/2}$ 轨道低约 1.50 eV，而同族其他元素的价层 p 轨道的能级分裂最多为 0.44 eV，所以虽然碳族元素的基态价电子构型均为 s^2p^2， Pb 的主要化合价是 2，而其他碳族元素的主要化合价是 4。不仅如此，由于 Pb 的价电子结构 ($6s^2 6p_{1/2}^2$) 类似于惰性气体的电子满壳层分布，所以 Pb 与其他元素的成键强度比第五周期的同族元素 Sn 弱很多。例如，Pb_2 的结合能为 0.98 eV，远小于 Sn_2 的结合能 1.95 eV；PbH 的结合能 (1.59 eV) 低于 SnH 的结合能 (3.2 eV)[9]。

Bi 的价电子结构 ($6s^2 6p_{1/2}^2 6p_{3/2}^1$) 与外层仅一个价电子的碱金属类似，所以 Bi 的第一电离能较低，且显示 +1 的化合价。对 Po 而言，$6p_{1/2}$ 轨道能量比 $6p_{3/2}$ 轨道低约 2.76 eV(见表 7.2)，$6p_{1/2}$ 轨道上占据的两个电子类似内层电子对，参与化学成键的主要是 $6p_{3/2}$ 轨道上的两个电子，所以 Po 的主要化合价为 2，而同族的 Se, Te 的化合价为 4。

对范德华分子 Rg_2 (Rg=Xe, Rn), RnXe 的研究表明旋–轨耦合效应使 Rn—Rn, Rn—Xe 键更加稳定，对 Xe—Xe 键的影响则可以忽略不计[6]。

3. 相对论效应与铅酸蓄电池

铅酸蓄电池是一种常见的蓄电池，广泛用于不间断电源、汽车牵引电源、控制开关等。铅酸蓄电池的正极二氧化铅板与负极铅板和电解液 (稀硫酸溶液) 反应，生成硫酸铅和水



$$Pb(s) + PbO_2(s) + 2H_2SO_4(aq) \longrightarrow 2PbSO_4(s) + 2H_2O(l) \tag{7.6}$$

式中，(l) 表示液态；(s) 表示固态；(aq) 表示水溶液。铅酸蓄电池将化学能转换为电能的能力由电动势 (electromotive force，EMF) E_{cell}^0 决定

$$\Delta G^0 = -nFE_{\text{cell}}^0 \tag{7.7}$$

式中，F 为 Faraday 常数；n 为反应中电子转移数；ΔG^0 表示电池化学反应 (7.6) 式中的自由能变化。Ahuja 等发现[7]，如果不考虑相对论效应，计算得到的铅酸蓄电池电动势只有 0.39V，远远低于实验值 (约为 2.107V)，考虑相对论效应以后，他们计算得到的电动势约为 2.13V，和实验值吻合得非常好，也就是说大约有 81% 的

电能来自于相对论效应的贡献。研究表明，由于铅中 6s 轨道的相对论收缩，二氧化铅中的 6s 态变得十分稳定，能带之间的间隔 (带隙) 显著降低[7]，他们发现对电动势相对论计算值贡献最大的是二氧化铅，其次是硫酸铅，再其次是金属铅中的旋–轨耦合效应。

7.2　过渡金属元素

过渡金属元素包括周期表中第 3~12 族等共十个族的元素，习惯上又以各族第一个元素将这十族元素分别称为：钪族、钛族、钒族、铬族、锰族、铁族、钴族、镍族、铜族、锌族等。它们的价层电子基态组态可用通式 $nd^m(n+1)s^{1,2}(n=3\sim5, m=1\sim10)$ 来表达 (见表 7.3)。本节将重点讨论第三过渡系金属 (除 La 外) 的 6s 轨道相对论收缩。La 和第四过渡系金属 Ac 又属于 f 族元素，关于它们的相对论效应读者请参见 7.3 节。其余第四过渡系金属的相对论效应我们将在超重元素一节 (7.4 节) 讨论。

表 7.3　第 5~12 族元素基态的 $ns(n=4\sim6)$ 轨道能量 ($-\varepsilon$，Hartree) 及 6s 轨道相对论收缩 ($\langle r_{6s}\rangle^{rel}/\langle r_{6s}\rangle^{nrel}$)[1]

	V	Cr	Mn	Fe	Co	Ni	Cu	Zn
电子组态	$3d^3 4s^2$	$3d^5 4s^1$	$3d^5 4s^2$	$3d^6 4s^2$	$3d^7 4s^2$	$3d^8 4s^2$	$3d^{10} 4s^1$	$3d^{10} 4s^2$
ε_{4s}	0.235	0.211	0.254	0.264	0.273	0.282	0.245	0.299
	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd
电子组态	$4d^4 5s^1$	$4d^5 5s^1$	$4d^5 5s^2$	$4d^7 5s^1$	$4d^8 5s^1$	$4d^{10}$	$4d^{10} 5s^1$	$4d^{10} 5s^2$
ε_{5s}	0.206	0.212	0.245	0.223	0.228		0.237	0.281
	Ta	W	Re	Os	Ir	Pt	Au	Hg
电子组态	$5d^3 6s^2$	$5d^4 6s^2$	$5d^5 6s^2$	$5d^6 6s^2$	$5d^7 6s^2$	$5d^9 6s^1$	$5d^{10} 6s^1$	$5d^{10} 6s^2$
ε_{6s}	0.254	0.267	0.278	0.289	0.299	0.284	0.292	0.328
6s 收缩	0.901	0.894	0.889	0.881	0.874	0.835	0.827	0.854

对第 4 族 (钛族) 的第三过渡系元素 Hf，由于相对论效应和原子壳结构膨胀效应互相抵消，Hf 的化学性质和第二过渡系的同族元素 Zr 区别并不大。从第 5 族 (钒族) 元素开始，第三过渡系金属的化学性质和同族第二过渡系金属的化学性质显著不同。由表 7.3 可以看出，第二、三过渡系金属价层电子填充顺序是不同的，第二过渡系金属优先填充 4d 轨道，而第三过渡系金属优先填充 6s 轨道，这是由于 6s 轨道的相对论收缩造成的。观察铜族的三个元素 Cu, Ag 和 Au，它们均有相同的价层电子结构 $nd^{10}(n+1)s^1(n=3\sim5)$，与 Cu, Ag 的价层 s 轨道相比，Au 的 6s 轨道能量 ε_{6s} 是最低的，所以 Au 的第一电离能要高于 Ag。同样由于 6s 轨道的

相对论收缩，其他第三过渡系金属 Ta, W, Re, Os, Ir, Pt 等的第一电离能也要高于同族第二过渡系金属 Nb, Mo, Tc, Ru, Rh, Pd 等。

表 7.3 还列出了 6s 轨道的相对论收缩 ($\langle r_{6s}\rangle^{rel}/\langle r_{6s}\rangle^{nrel}$)，很明显，6s 轨道的相对论收缩与 5d 轨道占据的电子数成正比—— 5d 轨道占据的电子数越多，6s 轨道的相对论收缩越显著。如表 7.3 所示，79 号元素金 (Au) 的基态 ($5d^{10}6s^1$) 有充满的 5d 轨道，所以 6s 轨道的相对论收缩也达到极大。研究表明，在核电荷数小于 100 的所有元素中，金的 6s 轨道相对论收缩也是最大的 (见图 7.1)。

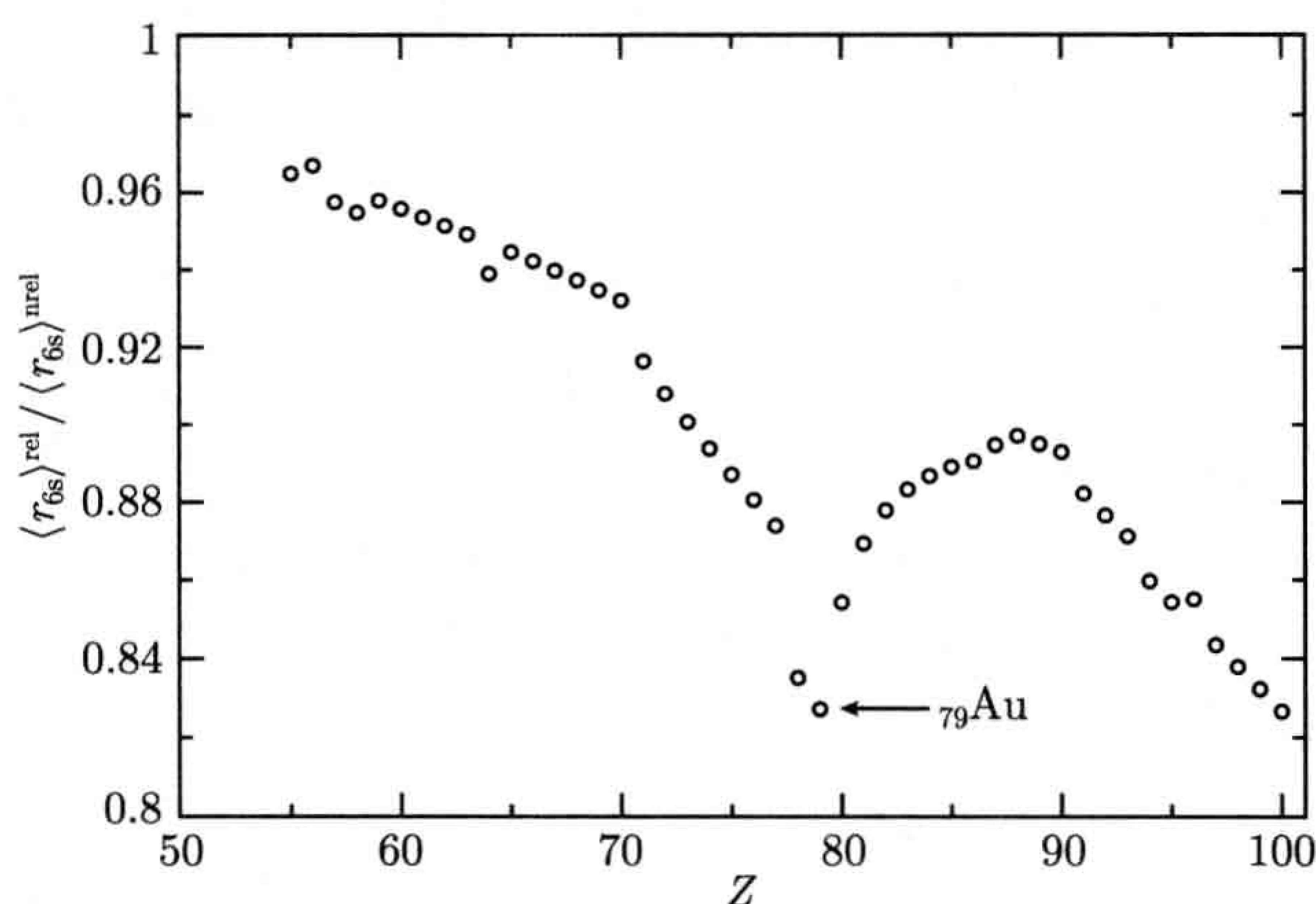


图 7.1 第三过渡系金属 6s 轨道的相对论收缩[1]

由于强烈的相对论效应，金的化学性质和同族的第二过渡系金属银 (Ag) 有很大不同，金的电离能更高、电子亲和性更强、能形成更短、更稳定的共价键等。表 7.4 列出了铜族元素二聚体 Cu_2, Ag_2 和 Au_2 的光谱结构常数，从表中可以看出，相对论计算结果和实验值吻合得相当好。二聚体 Au_2 比 Ag_2 的键长更短、结合能更高，三个二聚体的稳定顺序为 $Au_2 > Cu_2 > Ag_2$。另外，由于相对论收缩，金的 6s 轨道更靠近原子内部区域，金因此还显示了“卤素”的特征，如在半导体材料 CsAu 中 Au 的化合价为 −1。

表 7.4 Cu_2, Ag_2, Au_2 的基态 ($X^1\Sigma_g^+$) 光谱常数[8]

		R_e/Å	ω_e/cm^{-1}	D_e/eV
Cu_2	理论值	2.214	270.1	2.02
	实验值	2.2193	266.46	2.078±0.025
Ag_2	理论值	2.522	196.4	1.71
	实验值	2.5303	192.4	1.648±0.035
Au_2	理论值	2.472	191.5	2.31
	实验值	2.4715	191.05	2.320±0.004

注：1 cm $^{-1}$ = 1.9863×10^{-23} J。

对第 12 族 (锌族) 元素，相对论计算和实验研究表明第三过渡系金属汞的二聚体 Hg_2 比第二过渡系金属镉的二聚体 Cd_2 的键长更短、结合能更高[8]。与 Au_2, Ag_2 电子结构类似的二聚体 Hg_2^{2+}, Cd_2^{2+} 也符合同样的规律，即 Hg_2^{2+} 比 Cd_2^{2+} 更稳定。

由于过渡金属中普遍存在未充满的 d 轨道，所以过渡金属容易形成配合物，另外，6s 轨道的相对论收缩导致了 5d 的相对论膨胀 (间接相对论效应)，所以与 4d 轨道相比，5d 轨道的化学成键能力更强。例如，第二过渡系金属 Mo 最多与四个 Br 结合，形成 $MoBr_4$，而同族的第三过渡系金属 W 可以和五个或六个 Br 结合，形成 WBr_5 和 WBr_6。一般情况下，同族第二、三过渡系金属形成的离子型化合物，第三过渡系金属显示更高的化合价[9]。

相对论效应与过渡金属的颜色：当一束可见光 (能量范围 1.65~3.17 eV) 照射到金属表面时，金属会吸收特定波长范围的光，反射其他波长的光，人们肉眼看到的金属的颜色是由金属反射的光来决定的。而金属对特定波长光的吸收是由金属的电子能带结构决定的，这样我们就可以从金属的颜色变化中直接观察到相对论效应。以金为例，由于 6s 的相对论收缩和 5d 的相对论膨胀，相对论计算得到金的 Fermi 能级① (6s) 和 5d 能带的能量差仅为 2.38 eV (实验值约为 2.4 eV)，十分接近蓝光的能量范围 (2.50~2.76 eV)。当一束可见光照射到金表面时，金的 5d 能带电子会吸收蓝光跃迁至 Fermi 能级 (6s)，其余波长范围的可见光则被反射，所以我们看到的金是金色的。值得注意的是，对单个的金原子，由于 $5d^{10}6s^1 \rightarrow 5d^96s^2$ 的跃迁是宇称禁止的，所以单个的金原子并不呈现出金色。

与金同族的较轻元素银，其价层 4d, 5s 轨道的相对论效应要比金的 5d, 6s 轨道小得多，计算表明 4d, 5s 轨道的能级间隔约为 3.7 eV(与紫外光的能量接近)，所以银主要吸收可见光中的紫外光，呈现银白色。

7.3 镧系和锕系元素

镧系 (Ln) 和锕系 (An) 元素中存在未充满的 f 轨道，所以又称它们为 f 族元素。由于非直接相对论效应，这两族元素的价电子轨道 (d, f) 的能量升高、轨道半径增大。如图 7.2 和图 7.3 所示，与非相对论能量值相比，镧系原子基态的 4f 和 5d 轨道能量 (ε) 分别升高了 20%~35%, 11%~24%；轨道半径 ($\langle r \rangle$) 膨胀了 3%~8%, 4%~12% 左右。锕系元素的相对论效应更明显，其原子基态的 5f 和 6d 轨道能量分别升高了 40%~55% 和 20%~40%；轨道半径膨胀率分别为 7%~17% 和 8%~17%。f 族元素价层轨道能级的变化直接影响了价电子的分布，以 U 为例，由于非相对论计算得到的 5f 轨道能量过低，HF 计算发现 U 的基态应为 $5f^47s^2$，比实验观测到

① Fermi 能级：基态时电子的最高占据能级

的基态 $5f^36d^17s^2$ 低约 3.07 eV。考虑相对论效应后，DHF 计算得到的基态电子组态 $5f^36d^17s^2$ 与实验结果一致，激发态 $5f^47s^2$ 的能量比基态要高出约 2.04 eV[10]。

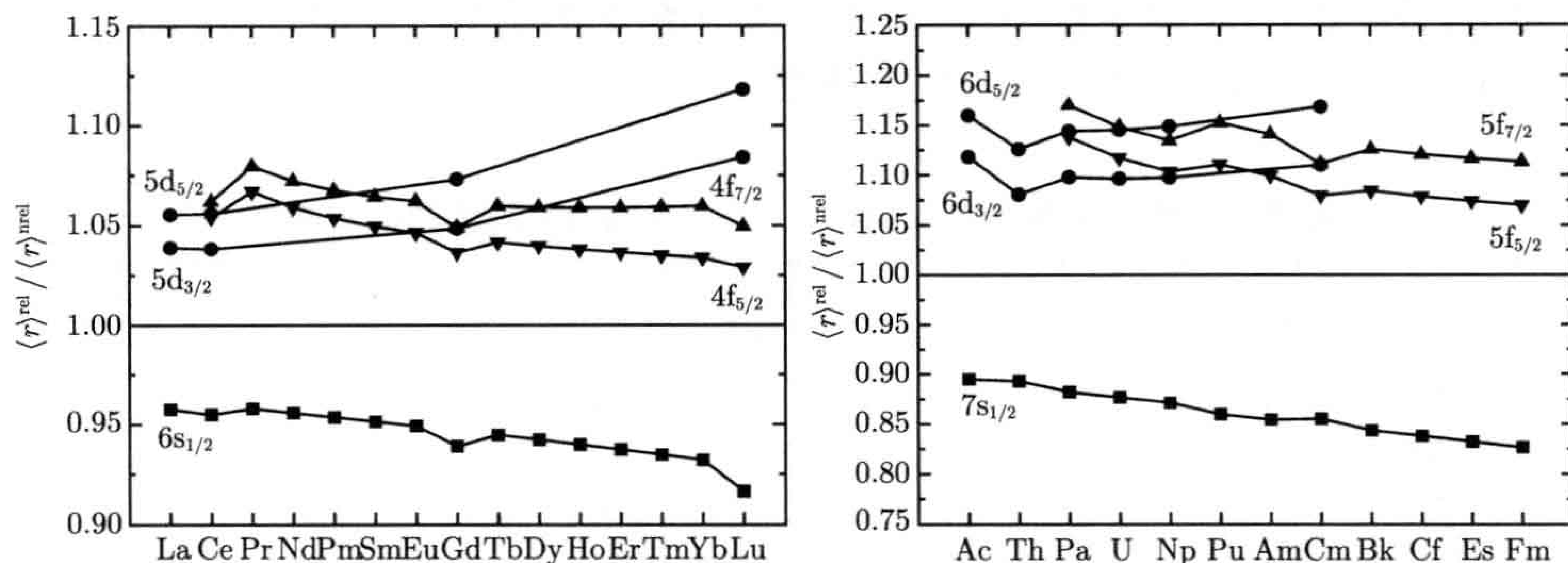


图 7.2 f 族元素基态价电子轨道半径的相对论收缩 $\langle r\rangle^{rel}/\langle r\rangle^{nrel}$[1]

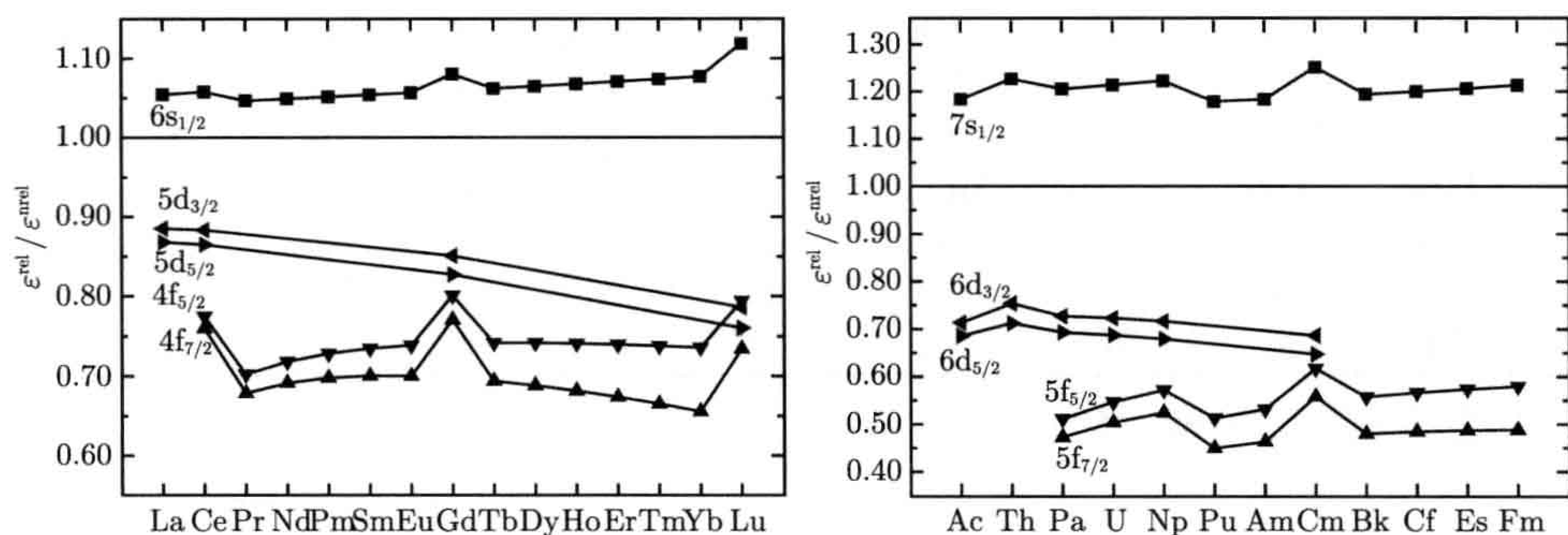


图 7.3 f 族元素基态价电子轨道能量的相对论收缩 $\varepsilon^{rel}/\varepsilon^{nrel}$[1]

价层 f 轨道强烈的非直接相对论效应也预示着当 f 轨道的电子布居数发生改变时，相对论效应对计算结果的影响也变大。表 7.5 和表 7.6 分别列出了镧系和锕系元素离子 (M^{n+}, M=Ln, An, $n = 0 \sim 4$) 的基态电子组态。容易发现，多数情况下 f 族元素的第三、四电离能的计算涉及价层 f 轨道电子布居数的改变。以锕系元素第四电离能计算 (IP_4, $M^{3+} \to M^{4+}$, M= Ac~Lr) 为例，由表 7.6 和图 7.4 可知，除 Ac(相对论效应对 IP_4 的贡献为 3%) 外，其余所有锕系元素的 IP_4 都涉及 5f 轨道电子布居数的改变，相对论效应对这些元素 IP_4 的贡献至少为 20%；再考虑锕系元素第三电离能 (IP_3)，对电离前后 5f 轨道电子布居数保持不变的 Ac, Th, Pa, Lr 等元素，相对论效应的贡献小于 15%，而对于其他元素，相对论效应对 IP_3 的贡献大于 30%。镧系元素的情况与锕系类似，当电离前后 4f 轨道电子布居数发生改变时 (见表 7.5)，相对论效应对电离能的贡献为 10%~23%；对电离前后 4f 轨道电子布居数保持不变的镧系元素，相对论效应的贡献约为 1%~6%，如图 7.4 所示。

由图 7.3 可知，随着原子序数的增高，f 轨道的自旋–轨道分裂也增大，从 Ce 到 Lu，其基态 4f 轨道的自旋–轨道能级分裂由 0.28 eV (Ce) 增加到 1.76 eV (Lu)；与 4f 轨道相比，5d 轨道的自旋–轨道能级分裂变化较小，分别为 0.14 eV (Ce), 0.17 eV (Lu)。锕系元素的自旋–轨道能级分裂更大，从锕系的第一个元素 Ac 到

表 7.5　镧系元素离子 (Ln^{n+}, $n = 0 \sim 4$) 的基态电子组态

	M		M^{1+}		M^{2+}		M^{3+}		M^{4+}	
La	d^1s^2	$^2D_{3/2}$	d^2	$^3F_2$	d^1	$^2D_{3/2}$		$^1S_0$	p^5	$^2P_{3/2}$
Ce	$f^1d^1s^2$	$^1G_4$	f^1d^2	$^4H_{7/2}$	f^2	$^3H_4$	f^1	$^2F_{5/2}$	p^6	$^1S_0$
Pr	f^3s^2	$^4I_{9/2}$	f^3s^1	$^5I_4$	f^3	$^4I_{9/2}$	f^2	$^3H_4$	f^1	$^2F_{5/2}$
Nd	f^4s^2	$^5I_4$	f^4s^1	$^6I_{7/2}$	f^4	$^5I_4$	f^3	$^4I_{9/2}$	f^2	$^3H_4$
Pm	f^5s^2	$^6H_{5/2}$	f^5s^1	$^7H_2$	f^5	$^6H_{5/2}$	f^4	$^5I_4$	f^3	$^4I_{9/2}$
Sm	f^6s^2	$^7F_0$	f^6s^1	$^8F_{1/2}$	f^6	$^7F_0$	f^5	$^6H_{5/2}$	f^4	$^5I_4$
Eu	f^7s^2	$^8S_{7/2}$	f^7s^1	$^9S_4$	f^7	$^8S_{7/2}$	f^6	$^7F_0$	f^5	$^6H_{5/2}$
Gd	$f^7d^1s^2$	$^9D_2$	$f^7d^1s^1$	$^{10}D_{5/2}$	f^7d^1	$^9D_2$	f^7	$^8S_{7/2}$	f^6	$^7F_0$
Tb	f^9s^2	$^6H_{15/2}$	f^9s^1	$^7H_8$	f^9	$^6H_{15/2}$	f^8	$^7F_6$	f^7	$^8S_{7/2}$
Dy	$f^{10}s^2$	$^5I_8$	$f^{10}s^1$	$^6I_{17/2}$	f^{10}	$^5I_8$	f^9	$^6H_{15/2}$	f^8	$^7F_6$
Ho	$f^{11}s^2$	$^4I_{15/2}$	$f^{11}s^1$	$^5I_8$	f^{11}	$^4I_{15/2}$	f^{10}	$^5I_8$	f^9	$^6H_{15/2}$
Er	$f^{12}s^2$	$^3H_6$	$f^{12}s^1$	$^4H_{13/2}$	f^{12}	$^3H_6$	f^{11}	$^4I_{15/2}$	f^{10}	$^5I_8$
Tm	$f^{13}s^2$	$^2F_{7/2}$	$f^{13}s^1$	$^3F_4$	f^{13}	$^2F_{7/2}$	f^{12}	$^3H_6$	f^{11}	$^4I_{15/2}$
Yb	$f^{14}s^2$	$^1S_0$	$f^{14}s^1$	$^2S_{1/2}$	f^{14}	$^1S_0$	f^{13}	$^2F_{7/2}$	f^{12}	$^3H_6$
Lu	$f^{14}d^1s^2$	$^2D_{3/2}$	$f^{14}s^2$	$^1S_0$	$f^{14}s^1$	$^2S_{1/2}$	f^{14}	$^1S_0$	f^{13}	$^2F_{7/2}$

注：f,d,s 分别表示镧系的 4f, 5d 和 6s 轨道。

表 7.6　锕系元素离子 (An^{n+}, $n = 0 \sim 4$) 的基态电子组态

	M		M^{1+}		M^{2+}		M^{3+}		M^{4+}	
Ac	d^1s^2	$^2D_{3/2}$	s^2	$^1S_0$	s^1	$^2S_{1/2}$		$^1S_0$	p^5	$^2P_{3/2}$
Th	d^2s^2	$^3F_2$	d^2s^1	$^4F_{3/2}$	f^1d^1	$^1G_4$	f^1	$^2F_{5/2}$	p^6	$^1S_0$
Pa	$f^2d^1s^2$	$^4K_{11/2}$	f^2s^2	$^3H_4$	f^2d^1	$^4I_{11/2}$	f^2	$^3H_4$	f^1	$^2F_{5/2}$
U	$f^3d^1s^2$	$^5K_6$	f^3s^2	$^4I_{9/2}$	f^4	$^5I_4$	f^3	$^4I_{9/2}$	f^2	$^3H_4$
Np	$f^4d^1s^2$	$^6L_{11/2}$	$f^4d^1s^1$	$^7L_5$	f^5	$^6H_{5/2}$	f^4	$^5I_4$	f^3	$^4I_{9/2}$
Pu	f^6s^2	$^7F_0$	f^6s^1	$^8F_{1/2}$	f^6	$^7F_0$	f^5	$^6H_{5/2}$	f^4	$^5I_4$
Am	f^7s^2	$^8S_{7/2}$	f^7s^1	$^9S_4$	f^7	$^8S_{7/2}$	f^6	$^7F_0$	f^5	$^6H_{5/2}$
Cm	$f^7d^1s^2$	$^9D_2$	f^7s^2	$^8S_{7/2}$	f^8	$^7F_6$	f^7	$^8S_{7/2}$	f^6	$^7F_0$
Bk	f^9s^2	$^6H_{15/2}$	f^9s^1	$^7H_8$	f^9	$^6H_{15/2}$	f^8	$^7F_6$	f^7	$^8S_{7/2}$
Cf	$f^{10}s^2$	$^5I_8$	$f^{10}s^1$	$^6I_{17/2}$	f^{10}	$^5I_8$	f^9	$^6H_{15/2}$	f^8	$^7F_6$
Es	$f^{11}s^2$	$^4I_{15/2}$	$f^{11}s^1$	$^5I_8$	f^{11}	$^4I_{15/2}$	f^{10}	$^5I_8$	f^9	$^6H_{15/2}$
Fm	$f^{12}s^2$	$^3H_6$	$f^{12}s^1$	$^4H_{13/2}$	f^{12}	$^3H_6$	f^{11}	$^4I_{15/2}$	f^{10}	$^5I_8$
Md	$f^{13}s^2$	$^2F_{7/2}$	$f^{13}s^1$	$^3F_4$	f^{13}	$^2F_{7/2}$	f^{12}	$^3H_6$	f^{11}	$^4I_{15/2}$
No	$f^{14}s^2$	$^1S_0$	$f^{14}s^1$	$^2S_{1/2}$	f^{14}	$^1S_0$	f^{13}	$^2F_{7/2}$	f^{12}	$^3H_6$
Lr	$f^{14}p^1s^2$	$^2P_{1/2}$	$f^{14}s^2$	$^1S_0$	$f^{14}s^1$	$^2S_{1/2}$	f^{14}	$^1S_0$	f^{13}	$^2F_{7/2}$

注：f,d,s 分别表示锕系的 5f, 6d 和 7s 轨道。

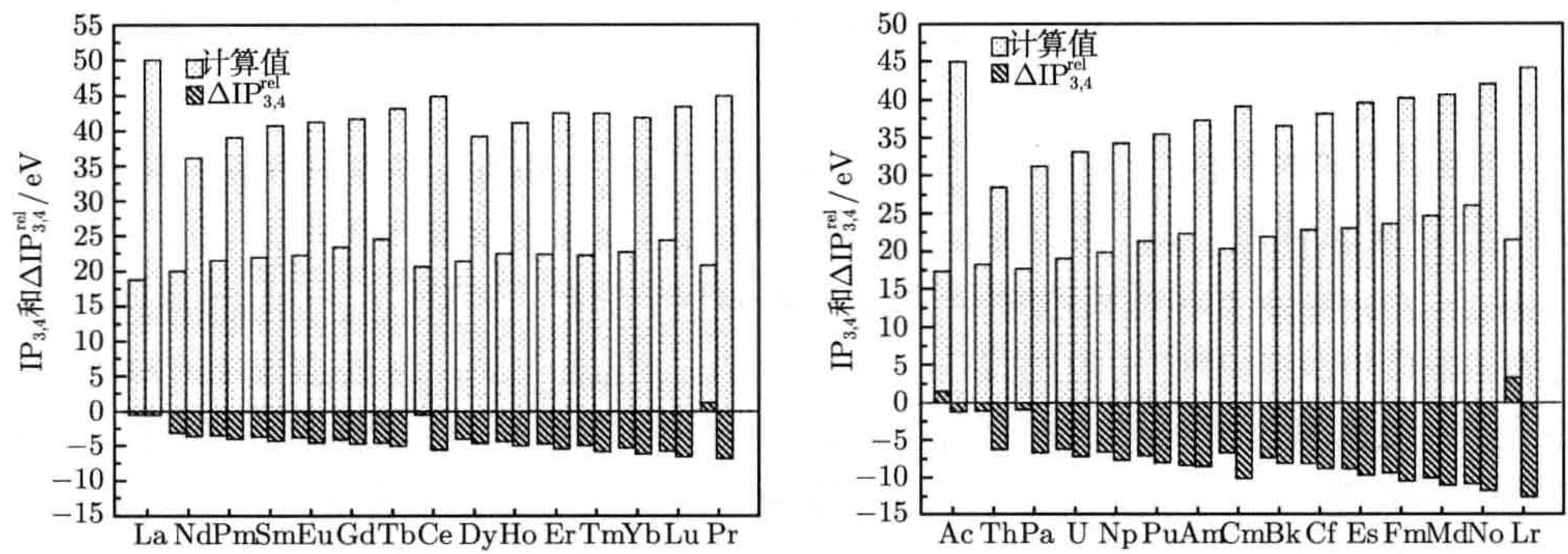


图 7.4 f 族元素第三、四电离能 (IP_3, IP_4) 计算值及相对论效应的贡献[10, 11] $\Delta\ IP^{rel}_{3,4}$

Cm，5d 轨道的能级分裂由 0.74 eV 上升到 1.07 eV；从 Pa 到 Fm，基态 5f 轨道的自旋–轨道能级分裂由 1.05 eV 增加到 2.49 eV(见图 7.3)。计算表明旋–轨耦合对镧系、锕系元素电离能的贡献，最高可达 2 eV 左右。

1. 镧 (锕) 系收缩

当原子序数增高时，f 族元素的原子 (离子) 半径会降低，称为镧 (锕) 系收缩。从镧系的第一个元素 La 到最后一个元素 Lu，原子半径收缩了大约 0.19 Å，三价阳离子半径收缩了约 0.187 Å；锕系元素的收缩比镧系元素更明显，从锕系的第一个元素 Ac 到最后一个元素 Lr，原子半径收缩了大约 0.3 Å。造成这种现象的原因是：① f 电子的屏蔽效应不完全；② 相对论效应。表 7.7 列出了从镧系的第一个元素 $La^{3+} \rightarrow Lu^{3+}$、锕系的第一个元素 $Ac^{3+} \rightarrow Lr^{3+}$ 最高占据轨道半径收缩 ($\Delta\langle r_{5p}\rangle$, $\Delta\langle r_{6p}\rangle$) 的 DHF 计算值、HF 计算值等。与实验推导的离子

表 7.7 镧系和锕系收缩的相对论 (rel) 及非相对论 (nrel) 计算值

f^0	f^{14}	收缩值/Å					
			nrel	rel	相对论贡献	实验值	文献
La^{3+}	Lu^{3+}	$\langle r_{5p}\rangle$①	0.188	0.204	0.016	0.171	[12]
LaH	LuH	R	0.115	0.113	−0.002	0.12	[13, 14]
LaF	LuF	R	0.097	0.111	0.014	0.11	[13, 14]
LaO	LuO	R	0.025	0.047	0.022	0.036	[13, 14]
LaF_3	LuF_3	R		0.198		0.2	[12]
Ac^{3+}	Lr^{3+}	$\langle r_{6p}\rangle$①	0.204	0.238	0.034	0.184	[12]
AcH	LrH	R	0.137	0.146	0.009		[13, 15]
AcF	LrF	R	0.084	0.114	0.030		[13, 15]
AcO	LrO	R	−0.012	0.071	0.083		[13, 15]
AcF_3	LrF_3	R	0.125	0.212	0.087		[12]

① 平均值。

半径收缩值相比，相对论效应对镧系和锕系元素离子半径收缩的贡献分别约为 9% 和 18%。

理论计算得到的轨道半径 $\langle r\rangle$ 并非可观测的物理量，人们一般也用分子键长的变化来表征镧 (锕) 系收缩。表 7.7 列出了 MF_3, MX (M=La, Lu, Ac, Lr; X=H, F, O) 等分子的键长收缩计算值和实验值。显然，不同分子的收缩幅度是不同的，LaO→LuO 的键长收缩最小，约为 0.04 Å(见表 7.7)，这是由于 LaO 分子中 La 的 4f 轨道参与了化学成键，而在 LuO 中满壳层的 4f 轨道类似内层轨道，并未参与化学成键[16]。对不同的分子，相对论效应对键长收缩的贡献也是不一样的，其中氧化物分子受其影响最大 (>60%)。另外，当镧系和锕系元素与相同配体成键时，相对论效应对锕系收缩的贡献大于相对论效应对镧系收缩的贡献。

2. 相对论效应和镧系化合物的发光

物质的发光是由于物质受到外界能量的激发，其电子从基态跃迁到激发态，当电子由激发态返回较低能级时，发射出不同波长的光。

镧系化合物的发光主要是由于三价金属阳离子 (Ln^{3+}) 中电子的 f − f 跃迁引起的，化合物发光的颜色由跃迁能级的能量差决定。研究表明，配体的改变对跃迁能级的影响不大[17]，下面我们就通过对 Eu^{3+} 的简单计算来说明相对论效应对镧系化合物发光的影响。

Eu^{3+} 的基态电子组态为 $^7F_0$: $[Xe]4f^6$，发光涉及的主要电子跃迁过程为 $^5D_0 \to {}^7F_n(n=0\sim4)$，即 $^5D_0$ 和 $^7F_n(n=0\sim4)$ 的能量差决定化合物发光的颜色。如果不包括相对论效应，激发态 ^5D 与基态 ^7F 的能级差约为 3.12 eV(非相对论 MCHF 计算结果)，考虑相关能效应后能级差降为 2.75 eV，接近蓝 (紫) 光的能量范围 (2.50~2.76 eV)，也就是说根据非相对论计算结果，含 Eu^{3+} 的化合物应该发蓝光，而实验现象是含 Eu^{3+} 的化合物发红光，二者并不一致。

包括相对论效应后，^5D 分裂成 5 个能级，^7F 分裂成 7 个能级，相对论 MCDHF 计算表明 $^5D_0$ 与 $^7F_n(n=0\sim4)$ 的能级差为 2.16~2.54 eV，考虑相关能效应后能级差约为 1.79~2.17 eV，接近可见光中红光区域 (1.65~2.10 eV)，即相对论计算结果显示 Eu^{3+} 镧系化合物发红光，符合实验观测结果。

7.4 超重元素

本节讨论元素周期表中核电荷数为 103~120 的元素的相对论效应。这些元素的质量显著大于其他元素，因而被称为超重元素，为便于讨论，我们统一用符号 “EZ” (Z 为原子序数) 来表示 (部分超重元素已有正式官方名称，见表 7.8)。超重元素全部为人工合成元素，具有放射性，且半衰期很短，如 E104 的半衰期小于

1h、E105 和 E106 的半衰期小于 1min 等，对原子序数大于 107 的超重元素，半衰期则小于 1s[18]。由于这些元素的质量很大，相对论效应较其他元素更为显著，主要表现在以下几个方面。

表 7.8 超重元素 (E104～E120) 基态的相对论和非相对论电子组态[18]

	相对论电子组态	J	非相对论电子组态	谱项
E104	$[Rn]5f^{14}6s^{2}6p^{6}6d^{2}7s^{2}$	2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{3}7s^{1}$	^{5}F
E105	$[Rn]5f^{14}6s^{2}6p^{6}6d^{3}7s^{2}$	3/2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{5}$	^{6}S
E106	$[Rn]5f^{14}6s^{2}6p^{6}6d^{4}7s^{2}$	0+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{5}7s^{1}$	^{7}S
E107	$[Rn]5f^{14}6s^{2}6p^{6}6d^{5}7s^{2}$	5/2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{6}7s^{1}$	^{6}D
E108	$[Rn]5f^{14}6s^{2}6p^{6}6d^{6}7s^{2}$	4+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{8}$	^{3}F
E109	$[Rn]5f^{14}6s^{2}6p^{6}6d^{7}7s^{2}$	9/2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{9}$	^{2}D
E110	$[Rn]5f^{14}6s^{2}6p^{6}6d^{8}7s^{2}$	3+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{1}0$	^{1}S
E111	$[Rn]5f^{14}6s^{2}6p^{6}6d^{9}7s^{2}$	5/2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{1}$	^{2}S
E112	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}$	0+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}$	^{1}S
E113	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{1}$	1/2−	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{1}$	^{2}P
E114	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{2}$	0+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{2}$	^{3}P
E115	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{3}$	3/2−	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{3}$	^{4}S
E116	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{4}$	2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{4}$	^{3}P
E117	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{5}$	3/2−	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{5}$	^{2}P
E118	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{6}$	0+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{6}$	^{1}S
E119	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{6}8s^{1}$	1/2+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{6}8s^{1}$	^{2}S
E120	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{6}8s^{2}$	0+	$[Rn]5f^{14}6s^{2}6p^{6}6d^{10}7s^{2}7p^{6}7d^{1}8s^{1}$	^{3}D

注：根据国际纯粹与应用化学联合会 (IUPAC) 的命名，E104～E112 的元素符号分别是 Rf, Db, Sg, Bh, Hs, Mt, Ds, Rg, Cn；E114 和 E116 的元素符号分别是 Pe 和 Lv。

1. 7(8)s 轨道的相对论收缩

受直接相对论效应的影响，相对论计算得到的 7s 轨道能量显著降低，以 E112 为例，相对论计算得到的 7s 轨道能量值比非相对论的结果低约 5.77 eV (见表 7.9)。7s 轨道能量的降低影响了超重元素 (E104～E111) 价层电子轨道的分布，从表 7.8 可以看出，相对论与非相对论计算得到的这些元素的基态电子组态是不一致的。例如，相对论计算发现 E104 的 $6d_{3/2}$ 轨道和 7s 轨道分别占据 2 个电子， 7s 轨道能量比 $6d_{3/2}$ 轨道低约 1.80 eV(见表 7.9)；而非相对论计算表明 E104 基态的 6d 轨道占据 3 个电子，7s 轨道仅占据一个电子，当电子数目增加时，电子将优先填充 6d 轨道。

表 7.10 列出了相对论效应对 MAu_n (M=Hg, E112；n=1, 9, 14) 分子结构的影响。从表中可以看出，考虑相对论效应后，对 E112Au, $E112Au_9$, $E112Au_{14}$ 等分子，E112— Au 键长分别缩短了约 0.27 Å, 0.37 Å 和 0.32 Å。相对论效应对以上三个化合物稳定性的影响并不一致，对 E112Au，分子解离能降低了 0.11 eV，分子

振动频率增加了约 4 cm^{-1}；对 $E112Au_9$ 和 $E112Au_{14}$，考虑相对论效应后分子解离能分别增加了 0.27 eV 和 0.01 eV。Pershina[19] 发现，E112 — Au 键的键能要比 Hg — Au 键的键能低 20~40 kJ·mol^{-1}，这是由于 7s 轨道的相对论收缩，E112 中 7s 轨道的成键能力要弱于同族第三过渡系金属 Hg 的 6s 轨道。

表 7.9　超重元素 (E104~E120) 价电子轨道能量的相对论及非相对论计算值 ($-\varepsilon$，Hartree)[18]

	相对论值						非相对论值			
	$6d_{3/2}$	$6d_{5/2}$	7s	$7p_{1/2}$	$7p_{3/2}$	8s	6d	7s	7p	7p
E104	0.207	0.188	0.274				0.318	0.193		
E105	0.253	0.224	0.297				0.365	0.201		
E106	0.297	0.257	0.319				0.412	0.208		
E107	0.341	0.289	0.340				0.459	0.214		
E108	0.384	0.320	0.362				0.505	0.220		
E109	0.428	0.351	0.383				0.552	0.225		
E110	0.472	0.382	0.405				0.600	0.230		
E111	0.517	0.412	0.428				0.648	0.234		
E112	0.563	0.442	0.451				0.696	0.238		
E113	0.762	0.621	0.588	0.265	0.151		0.914	0.327	0.184	
E114	0.956	0.792	0.722	0.356	0.184		1.129	0.413	0.227	
E115	1.153	0.964	0.858	0.446	0.216		0.346	0.500	0.268	
E116	1.352	1.138	0.998	0.539	0.246		1.567	0.589	0.310	
E117	1.556	1.314	1.144	0.637	0.276		1.792	0.680	0.352	
E118	1.764	1.493	1.296	0.739	0.306		2.021	0.774	0.394	
E119	2.108	1.809	1.562	0.955	0.435	0.159	2.417	0.993	0.570	0.110
E120	2.455	2.126	1.830	1.172	0.560	0.192	2.813	1.201	0.735	0.138

注：由 IUPAC 命名的超重元素符号见表 7.8。

表 7.10　MAu_n (M=Hg, E112；n = 1, 9, 14) 的理论光谱常数[19]

		R_e/Å	ω_e/cm^{-1}	D_e/eV
HgAu	非相对论值	2.879	75.89	0.36
	相对论值	2.667	102.96	0.49
E112Au	非相对论值	2.990	77.19	0.47
	相对论值	2.725	81.74	0.36
$HgAu_9$	非相对论值	2.434		0.41
	相对论值	2.011		0.85
$E112Au_9$	非相对论值	2.593		0.52
	相对论值	2.223		0.79
$HgAu_{14}$	非相对论值	2.910		0.64
	相对论值	2.646		0.86
$E112Au_{14}$	非相对论值	3.069		0.70
	相对论值	2.752		0.71

如表 7.8 和表 7.9 所示，8s 轨道是 E119 和 E120 的价电子轨道，该轨道的相对论收缩效应也十分明显。研究表明，由于 8s 轨道的相对论收缩，E119 的电子亲和能 (EA) 为 662 meV[20]，远高于同族的较轻元素 Fr (491 meV)，E119 也是碱金属中电子亲和能最大的元素。对 E120，同样由于 8s 轨道的相对论收缩，相对论计算得到基态 E120 的 8s 轨道占据两个电子，而非相对论计算的结果是这两个电子分别占据 7d 和 8s 轨道。

观察表 7.9 我们发现，7s 轨道的相对论收缩大于 7p 轨道的相对论收缩，包括相对论效应后计算所得的 7s — 7p 的能量间隔变大，不利于形成 sp 杂化轨道，从而阻扰了 7s 电子参与化学成键。所以对 E113~E118 等超重元素，7s 轨道显示了内层成对电子效应，这些元素的氧化态均要低于同族的较轻元素。例如，对 E113，+1 的氧化态要比 +3 的氧化态更重要；对 E114，主要氧化态是 +2，而不是 +4 等。

2. 6d 轨道的相对论膨胀

对 E104~E111 等超重元素 (又称第四过渡系金属)，存在着空的 6d 轨道，膨胀后的 6d 轨道又易于杂化，所以这些元素与其他元素配位时容易形成不同的氧化态。

3. 7p 轨道强烈的旋-轨耦合效应

如表 7.9 所示，超重元素 (E113~E120) 基态的 $7p_{1/2}$ 轨道和 $7p_{3/2}$ 轨道的能级分裂很大，二者的能量间隔约为 3.13~16.65 eV。与非相对论的 7p 轨道比较，$7p_{3/2}$ 轨道能量升高 (间接相对论效应)，$7p_{1/2}$ 轨道能量降低 (直接相对论效应)——计算表明，$7p_{1/2}$ 轨道的能量要低于同族较轻元素的 $6p_{1/2}$ 轨道的能量。对双原子分子 MH(M=E113~E118) 的研究表明，由于 E113— H 中仅有能量很低的 $7p_{1/2}$ 轨道参与化学成键，相对论计算得到的 E113— H 键长比非相对论键长短约 0.2 Å，对其他双原子分子，$7p_{1/2}$ 轨道和 $7p_{3/2}$ 轨道均参与化学成键，相对论计算得到的 M—H (M= E114~E118) 键长要大于非相对论的结果[21]。

4. 量子电动力学 (QED) 效应

QED 效应是相对论效应的高阶校正，包括电子自能、电磁场真空极化等 (见附录 D)。Pyykkö 等发现，考虑 QED 效应后，超重元素 (E111~E119) 的电离能上升了 0.5%[22]；对 E118，QED 效应使 E118 的电子亲和能降低了 9%[23]。

参 考 文 献

[1] Desclaux J P. Relativistic Dirac-Fock expectation values for atoms with $Z = 1$ to $Z = 120$. At Data Nucl Data Tables, 1973, 12: 311~406.

[2] Fricke B, Waber J T. Atomic and ionic radii of superheavy elements. J Chem Phys, 1972, 56: 3246~3248.

[3] Martin W C, Musgrove A, Kotochigova A, et al. Ground levels and ionization energies for the neutral atoms. http://www.nist.gov/pml/data/ion_energy.cfm [2012-06-20].

[4] Pyykkö P. Relativistic effects in structural chemistry. Chem Rev, 1988, 88: 563~594.

[5] Pyykkö P, Desclaux J P. Dirac-Fock one-electron calculations the molecules BH, AlH, GaH, InH, and TlH. Chem Phys Lett, 1976, 42: 545~549.

[6] Runeberg R, Pyykkö P. Relativistic pseudopotential calculations on Xe_2, RnXe, and Rn_2: The van der Waals properties of radon. Int J Quan Chem, 1998, 66: 131~140.

[7] Ahuja R, Blomqvist A, Larsson P, et al. Relativity and the lead-acid battery. Phys Rev Lett, 2011, 106: 018301-1~018301-4.

[8] Peterson K A, Puzzarini C. Systematically convergent basis sets for transition metals II Pseudopotential-based correlation consistent basis sets for the group 11 (Cu, Ag, Au) and 12 (Zn, Cd, Hg) elements. Theor Chem Acc, 2005, 114: 283~296.

[9] Pyykkö P. Relativistic quantum chemistry. Adv Quant Chem, 1978, 11: 353~409.

[10] Cao X, Dolg M. Relativistic energy-consistent *ab initio* pseudopotentials as tools for quantum chemical investigations of actinide systems. Coord Chem Rev, 2006, 250: 900~912.

[11] Dolg M, Cao X. The relativistic energy-consistent *ab initio* pseudopotential approach and its application to lanthanide and actinide compounds. // Hirao K, Ishikawa Y. Recent Advances in Computational Chemistry. Vol 6. New Jersey: World Scientific, 2004, 1~35.

[12] Dolg M, Stoll H. Electronic structure calculations for molecules containing lanthanide atoms. // Eyring L. Handbook of Chemistry and Physics of Rare Earths. Vol 22. Amsterdam: Elsevier Science B V, 1996, 607~729.

[13] Küchle W, Dolg M, Stoll H. *Ab initio* study of the lanthanide and actinide contraction. J Phys Chem, 1997, 101: 7128~7133.

[14] Cao X. New basis sets for lanthanide and actinide energy-consistent small-core pseudopotentials. J Chin Chem Soc, 2003, 50: 665~676.

[15] Cao X, Dolg M. Segmented contraction scheme for small-core actinide pseudopotential basis sets. J Molec Struct (Theochem), 2004, 673: 203~209.

[16] 曹晓燕, 刘文剑, Dolg M. 双原子镧系化合物分子结构. 中国科学 B，2001, 31: 481~486.

[17] Cotton S. Lanthanide and Actinide Chemistry. Chichester: John Wiley & Sons, Ltd, 2006.

[18] Schwerdtfeger P, Seth M. Relativistic effects of the superheavy elements. // Schleyer P v R, Allinger N L, Clark T, et al. Encyclopedia of Computational Chemistry. Chichester: Wiley, 1998: 2480~2499.

[19] Pershina V. Electronic structure and chemistry of the superheavy elements. // Barysz

M, Ishikawa Y. Relativistic methods for chemists: Challenges and advances in computational chemistry and physics 10. Dordrecht: Springer Science+Business Media B V, 2010: 451~520.

[20] Landau A, Eliav E, Ishikawa Y, et al. Benchmark calculations of electron affinities of the alkali atoms sodium to eka-francium (element 119). J Chem Phys, 2001, 115: 2389~2392.

[21] Pershina V. Electronic structure and chemical properties of superheavy elements. Russ Chem Rev, 2009, 78: 1153~1171.

[22] Pyykkö P, Tokman M, Labzowsky L. Estimated valence-level Lamb shifts for group 1 and group 11 metal atoms. Phys Rev A, 1998, 57: R689~R692.

[23] Pyykkö P. The physics behind chemistry and the periodic table. Chem Rev, 2012, 112: 379~394.

附录A　算 符 代 数

在矢量空间中，规定一个具体的对应关系，使空间中的某些矢量与另一些矢量相对应，如使 φ 与 ψ 相对应，此对应关系称为算符，用 $\hat{O}$ 表示

$$\psi = \hat{O}\varphi \tag{A.1}$$

算符是物理学中的一个重要概念，下面列出本书中出现的几种算符名称的定义，其详细特征可参阅有关物理书籍。

1. 线性算符和反线性算符

一个算符 $\hat{O}$ 如果满足

$$\begin{aligned} \hat{O}(\varphi+\psi) &= \hat{O}\varphi + \hat{O}\psi \\ \hat{O}c\varphi &= c\hat{O}\varphi \end{aligned} \tag{A.2}$$

式中，c 是任意常数，则 $\hat{O}$ 称为线性算符。如果算符 $\hat{O}$ 满足

$$\begin{aligned} \hat{O}(\varphi+\psi) &= \hat{O}\varphi + \hat{O}\psi \\ \hat{O}c\varphi &= c^*\hat{O}\varphi \end{aligned} \tag{A.3}$$

式中，c^* 是 c 的复共轭，则 $\hat{O}$ 称为反线性算符。

2. 厄米算符和反厄米算符

如果一个算符 $\hat{O}$ 和它的转置复共轭算符 $\hat{O}^\dagger$ (也称 $\hat{O}$ 的伴算符) 相等，即

$$\hat{O} = \hat{O}^\dagger \tag{A.4}$$

则 $\hat{O}$ 称为厄米算符。如果

$$\hat{O} = -\hat{O}^\dagger \tag{A.5}$$

则 $\hat{O}$ 称为反厄米算符。

一个关于厄米算符的重要定理是：$\hat{O}$ 为厄米算符的充分必要条件是对算符定义域的任意 φ，有

$$\int \varphi^*\hat{O}\varphi \mathrm{d}\tau = \text{实数} \tag{A.6}$$

由上可以得到推论：表示可观测物理量的算符必为厄米算符。

3. 酉算符和反酉算符

如果一个算符 $\hat{U}$ 和它的转置复共轭算符 $\hat{U}^\dagger$ 满足

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = 1 \tag{A.7}$$

则 $\hat{U}$ 称为酉算符，英文名称为 unitary operator，也译为幺正算符。

如果一个反线性算符 $\hat{O}$ 满足

$$\langle\hat{O}\varphi|\hat{O}\psi\rangle = \langle\varphi|\psi\rangle^* = \langle\psi|\varphi\rangle \tag{A.8}$$

则 $\hat{O}$ 称为反酉算符。英文名称为 antiunitary operator，也译为反幺正算符。容易证明酉算符和反酉算符的积仍为反酉算符，两个反酉算符相乘得酉算符。

酉算符在讨论两组基矢的关系时起到重要作用，下面给出两条有关定理。

(1) 在矢量空间中，如果将酉算符 $\hat{U}$ 作用到一组基矢 $\{\varphi_i\}$，则得到的一组新矢量 $\{\hat{U}\varphi_i\}$ 也是基矢量。

(2) 若 $\{\varphi_i\}$ 和 $\{\psi_i\}$ 是同一矢量空间的两组基矢，则必能找到一个酉算符 $\hat{U}$，使 $\psi_i = \hat{U}\varphi_i$。

4. 投影算符

对空间中的一组基矢 $\{\varphi_i\}$，采用 Dirac 符号①，定义投向子空间 φ_i 的投影算符为

$$\hat{P}_i = |\varphi_i\rangle\langle\varphi_i| \tag{A.9}$$

投影算符的重要性质是它的幂等性，即 $\hat{P}_i^2 = \hat{P}_i$。

我们也可以定义投向整个空间的投影算符

$$\hat{P} = \sum_i |\varphi_i\rangle\langle\varphi_i| \tag{A.10}$$

既然是向整个空间投影，则矢量投影后不会发生任何变化，由此我们得出

$$\hat{P} = \sum_i |\varphi_i\rangle\langle\varphi_i| \equiv 1 \tag{A.11}$$

这是一个非常有用的关系，称为完全性关系 (也称为投影算符的封闭性)。此关系在计算中很有用，如计算两个函数的积分时，可以插入完全性关系以简化计算

$$\langle\varPsi|\varPhi\rangle = \sum_i \langle\varPsi|\varphi_i\rangle\langle\varphi_i|\varPhi\rangle \tag{A.12}$$

最后介绍一个常用的算符等式。

假设 $\hat{A}$ 和 $\hat{B}$ 为两个线性算符，且互不对易，则 $\hat{A}$ 和 $\hat{B}$ 的多重对易式 $[\hat{A}^{(n)}, \hat{B}]$ $(n = 0, 1, 2, \cdots)$ 定义为

① Dirac 定义矢量空间的态矢量为右矢 (ket): $|\varphi\rangle$；每一个右矢的转置复共轭定义为其左矢 (bra): $\langle\varphi|$

$$
\begin{aligned}
&\left[\hat{A}^{(0)}, \hat{B}\right] = B \\
&\left[\hat{A}^{(1)}, \hat{B}\right] = \left[\hat{A}, \hat{B}\right] \\
&\left[\hat{A}^{(2)}, \hat{B}\right] = \left[\hat{A}, \left[\hat{A}^{(1)}, \hat{B}\right]\right] \\
&\left[\hat{A}^{(3)}, \hat{B}\right] = \left[\hat{A}, \left[\hat{A}^{(2)}, \hat{B}\right]\right] \\
&\qquad\cdots \\
&\left[\hat{A}^{(n)}, \hat{B}\right] = \left[\hat{A}, \left[\hat{A}^{(n-1)}, \hat{B}\right]\right]
\end{aligned}
\tag{A.13}
$$

为和算符 $\hat{A}$ 的 n 次幂相区别，用来表示多重对易的上标 n 放在圆括号内。

定义算符函数 $\hat{F}(\lambda)$ 为

$$
\hat{F}(\lambda) = \mathrm{e}^{\lambda\hat{A}}\hat{B}\mathrm{e}^{-\lambda\hat{A}} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\left(\frac{\mathrm{d}^n F}{\mathrm{d}\lambda^n}\right)_{\lambda=0} \tag{A.14}
$$

式中，第二个等号为 $\hat{F}(\lambda)$ 的 Taylor 展开式。接下来求 $\hat{F}(\lambda)$ 的各阶导数在 $\lambda = 0$ 的值。

当 $n = 1$ 时

$$
\begin{aligned}
\left(\frac{\mathrm{d}\hat{F}(\lambda)}{\mathrm{d}\lambda}\right)_{\lambda=0} &= \left(\mathrm{e}^{\lambda\hat{A}}\hat{A}\hat{B}\mathrm{e}^{-\lambda\hat{A}} - \mathrm{e}^{\lambda\hat{A}}\hat{B}\hat{A}\mathrm{e}^{-\lambda\hat{A}}\right)_{\lambda=0} \\
&= \left(\mathrm{e}^{\lambda\hat{A}}\left[\hat{A}, \hat{B}\right]\mathrm{e}^{-\lambda\hat{A}}\right)_{\lambda=0} \\
&= \left[\hat{A}, \hat{B}\right]
\end{aligned}
\tag{A.15}
$$

当 $n = 2$ 时

$$
\begin{aligned}
\left(\frac{\mathrm{d}^2\hat{F}(\lambda)}{\mathrm{d}\lambda^2}\right)_{\lambda=0} &= \left(\mathrm{e}^{\lambda\hat{A}}\left[\hat{A}, \left[\hat{A}, \hat{B}\right]\right]\mathrm{e}^{-\lambda\hat{A}}\right)_{\lambda=0} \\
&= \left[\hat{A}, \left[\hat{A}, \hat{B}\right]\right] = \left[\hat{A}^{(2)}, \hat{B}\right]
\end{aligned}
\tag{A.16}
$$

上式最后一个等号利用了多重对易式的定义 (A.13) 式。以此类推，不难得到 $\hat{F}(\lambda)$ 的任意阶导数 $\lambda = 0$ 的值

$$
\begin{aligned}
\left(\frac{\mathrm{d}^n F}{\mathrm{d}\lambda^n}\right)_{\lambda=0} &= \left[\hat{A}, \left[\hat{A}, \left[\hat{A}, \cdots, \left[\hat{A}, \hat{B}\right]\right]\right]\right] \\
&= \left[\hat{A}^{(n)}, \hat{B}\right]
\end{aligned}
\tag{A.17}
$$

将结果代入 (A.14) 式，并令 $\lambda = 1$，得

$$
\mathrm{e}^{\hat{A}}\hat{B}\mathrm{e}^{-\hat{A}} = \sum_{n=0}^{\infty} \frac{1}{n!}\left[\hat{A}^{(n)}, \hat{B}\right] \tag{A.18}
$$

附录B　矢量分析

本附录列出矢量分析的常用公式，其详细证明可参阅有关的数学书籍。

B.1　矢量关系

1. 三矢量的混合积

$$\boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})=\boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a}) \tag{B.1}$$

2. 三矢量的矢积

$$\boldsymbol{c}\times(\boldsymbol{a}\times\boldsymbol{b})=(\boldsymbol{c}\cdot\boldsymbol{b})\,\boldsymbol{a}-(\boldsymbol{c}\cdot\boldsymbol{a})\,\boldsymbol{b} \tag{B.2}$$

3. Pauli 矢量规则

$$(\boldsymbol{\sigma}\cdot\boldsymbol{a})\,(\boldsymbol{\sigma}\cdot\boldsymbol{b})=\boldsymbol{I}_2\boldsymbol{a}\cdot\boldsymbol{b}+\mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{a}\times\boldsymbol{b}) \tag{B.3}$$

式中，$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 代表任意矢量；$\boldsymbol{\sigma}$ 为 Pauli 矩阵。

B.2　Nabla 算符 ∇

在直角坐标系 (x,y,z) 和球坐标系 (r,θ,φ) 中 ∇ 算符分别定义为

$$\nabla=\hat{\boldsymbol{e}}_x\frac{\partial}{\partial x}+\hat{\boldsymbol{e}}_y\frac{\partial}{\partial y}+\hat{\boldsymbol{e}}_z\frac{\partial}{\partial z} \tag{B.4}$$

$$\nabla=\hat{\boldsymbol{e}}_r\frac{\partial}{\partial r}+\hat{\boldsymbol{e}}_\theta\frac{1}{r}\frac{\partial}{\partial\theta}+\hat{\boldsymbol{e}}_\varphi\frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi} \tag{B.5}$$

∇^2 算符又称 Laplace 算符，用 ∇ 表示。以下列出了 ∇ 算符的运算公式。公式中 ϕ,ψ 代表任意标量，$\boldsymbol{a},\boldsymbol{b}$ 代表任意矢量。

$$\nabla\,(\phi\psi)=\phi\nabla\psi+\psi\nabla\phi \tag{B.6}$$

$$\nabla\cdot(\phi\boldsymbol{a})=(\nabla\phi)\cdot\boldsymbol{a}+\phi\nabla\cdot\boldsymbol{a} \tag{B.7}$$

$$\nabla\times(\phi\boldsymbol{a})=(\nabla\phi)\times\boldsymbol{a}+\phi\nabla\times\boldsymbol{a} \tag{B.8}$$

$$\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) = (\nabla \times \boldsymbol{a}) \cdot \boldsymbol{b} - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b}) \tag{B.9}$$

$$\nabla \times (\boldsymbol{a} \times \boldsymbol{b}) = (\boldsymbol{b} \cdot \nabla)\,\boldsymbol{a} + (\nabla \cdot \boldsymbol{b})\,\boldsymbol{a} - (\boldsymbol{a} \cdot \nabla)\,\boldsymbol{b} - (\nabla \cdot \boldsymbol{a})\,\boldsymbol{b} \tag{B.10}$$

$$\nabla (\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \times (\nabla \times \boldsymbol{b}) + (\boldsymbol{a} \cdot \nabla)\,\boldsymbol{b} + \boldsymbol{b} \times (\nabla \times \boldsymbol{a}) + (\boldsymbol{b} \cdot \nabla)\,\boldsymbol{a} \tag{B.11}$$

$$\nabla \times (\nabla \times \boldsymbol{a}) = \nabla (\nabla \cdot \boldsymbol{a}) - \nabla^2 \boldsymbol{a} \tag{B.12}$$

$$\nabla \cdot (\boldsymbol{a}\boldsymbol{b}) = (\nabla \cdot \boldsymbol{a})\,\boldsymbol{b} + (\boldsymbol{a} \cdot \nabla)\,\boldsymbol{b} \tag{B.13}$$

$$\boldsymbol{r} = \hat{\boldsymbol{e}}_x x + \hat{\boldsymbol{e}}_y y + \hat{\boldsymbol{e}}_z z \quad \Longrightarrow \quad \nabla \cdot \boldsymbol{r} = 3, \quad \nabla \times \boldsymbol{r} = 0 \tag{B.14}$$

在坐标表象下，动量算符可用 Nabla 算符表示为 $\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla$，动量算符与任一标量 ϕ 的对易关系可写为

$$[\hat{\boldsymbol{p}}, \phi] = -\mathrm{i}\hbar\nabla\phi \tag{B.15}$$

这是一个很有用的对易关系。

B.3　散度、旋度和梯度

利用 Nabla 算符 ∇，可以把矢量场 $\boldsymbol{a}$ 的散度和旋度、标量场 ϕ 的梯度表示为

$$\mathrm{div}\boldsymbol{a} = \nabla \cdot \boldsymbol{a} \tag{B.16}$$

$$\mathrm{rot}\boldsymbol{a} = \nabla \times \boldsymbol{a} \tag{B.17}$$

$$\mathrm{grad}\phi = \nabla\phi \tag{B.18}$$

关于散度和旋度，有如下重要定理

1. 标量场的梯度必为无旋场。

$$\nabla \times (\nabla\phi) \equiv 0 \tag{B.19}$$

2. 矢量场的旋度必为无源场。

$$\nabla \cdot (\nabla \times \boldsymbol{a}) = 0 \tag{B.20}$$

3. 无旋场必可表示为标量场的梯度。

$$\text{若} \quad \nabla \times \boldsymbol{a} = 0, \qquad \text{则} \quad \boldsymbol{a} = \nabla\phi \tag{B.21}$$

4. 无源场必可表示为另一矢量的旋度。

$$\text{若} \quad \nabla \cdot \boldsymbol{a} = 0, \qquad \text{则} \quad \boldsymbol{a} = \nabla \times \boldsymbol{b} \tag{B.22}$$

附录 C　原子单位制

原子单位制 (atomic unit，a.u.) 是一套广泛应用于量子化学的单位制。在这种单位制中，以下四个物理学常量的数值均为 1：

- 电子的静止质量 (m_e)
- 电子的电荷 (e)
- 约化 Planck 常量 ($\hbar$)
- Coulomb 常数 $\left(\dfrac{1}{4\pi\epsilon_0}\right)$

采取以上定义后，可以导出其他基本单位。表 C.1 列出了原子单位制中的常见单位 (物理量) 及其和国际单位制的关系。

表 C.1　常用单位及物理常量简表

名称	符号	原子单位	国际单位
电子的静止质量	m_e	1	9.1094×10^{-31} kg
Bohr 半径	$a_0=\hbar/(m_e c\alpha)$	1	5.2918×10^{-11} m
基本电荷	e	1	1.6022×10^{-19} C
约化 Planck 常量	$\hbar=h/(2\pi)$	1	1.0546×10^{-34} J·s
Hartree 能量	$E_h=e^2/a_0$	1	4.3597×10^{-18} J
Coulomb 常数	$1/(4\pi\epsilon_0)$	1	8.9876×10^{9} $C^{-2}Nm^2$
真空介电常量	ϵ_0	1	8.8542×10^{-12} $J^{-1}C^2\cdot m^{-1}$
Rydberg 常数	$R_\infty=m_e e^4/2\hbar^2$	0.5	2.1799×10^{-18} J
时间	$\hbar/E_h$	1	2.4189×10^{-17} s
速度	$a_0E_h/\hbar$	1	$2.1877\times10^{6}m\cdot s^{-1}$
力	E_h/a_0	1	8.2387×10^{-8} N
真空光速	c	137.0360	$2.9979\times10^{8}m\cdot s^{-1}$
精细结构常数	$\alpha=e^2/(4\pi\epsilon_0\hbar)c$	1/137.0360	1/137.0360
质子静质量	m_p	1 836.1527	1.6726×10^{-27} kg
中子静质量	m_n	1 838.6837	1.6749×10^{-27} kg

注：数据 (准确到小数点后四位) 来自 http://physics.nist.gov/cuu/Constants/index.html。

在原子单位制中，电荷为一个原子单位的两个质点，相隔一个原子单位的距离时，势能 e^2/a_0 定义为一个原子单位的能量。这个能量单位称为 Hartree，选取 Hartree 作为能量单位的原子单位制又称为 Hartree 单位制，表 C.2 列出了常见能量单位的换算关系。

量子化学文献中也有定义核质量 ∞ 时的 Rydberg 常数作为一个原子单位的能量，此种原子单位制称为 Rydberg 单位制。本书采用的均为 Hartree 单位制。

表 C.2　常用能量单位的换算关系

	J	kJ·mol^{-1}	erg	kcal·mol^{-1}	eV	Hartree	cm^{-1}
1 J	1	6.0229×10^{20}	10^{7}	1.4395×10^{20}	6.2420×10^{18}	2.2937×10^{17}	5.0345×10^{22}
1 kJ·mol^{-1}	1.6603×10^{-21}	1	1.6603×10^{-14}	2.3900×10^{-1}	1.0363×10^{-2}	3.8088×10^{-4}	8.3593×10^{1}
1 erg	10^{-7}	6.0229×10^{13}	1	1.4394×10^{13}	6.2418×10^{11}	2.2937×10^{10}	5.0345×10^{15}
1 kcal·mol^{-1}	6.9467×10^{-21}	4.184	6.9473×10^{-14}	1	4.3363×10^{-2}	1.5936×10^{-3}	3.4976×10^{2}
1 eV	1.6022×10^{-19}	9.6485×10	1.6021×10^{-12}	2.3061×10	1	3.6750×10^{-2}	8.0657×10^{3}
1 Hartree	4.3597×10^{-18}	2.6255×10^{3}	4.3597×10^{-11}	6.2750×10^{2}	27.2107	1	2.1947×10^{5}
1 cm^{-1}	1.9863×10^{-23}	1.1963×10^{-2}	1.9863×10^{-16}	2.8591×10^{-3}	1.2398×10^{-4}	4.5563×10^{-6}	1

注: $1Å= 10^{-10}$ m$= 10^{-8}$ cm $= 10^{-4}\mu$m $= 0.1$ nm。

附录D　QED 效应

光与物质相互作用的理论称为量子电动力学① (quantum electrodynamics, QED)。根据 Fermi 的表述，假设整个宇宙装在一个盒子里，宇宙中所有原子都在这个盒子里，这个盒子可以看成有一些本征态，这些态可以用量子力学谐振子以及这些振子与物质之间的耦合来描述。这些谐振子具有能量 $\left(n+\dfrac{1}{2}\right)\hbar\omega(n=0,1,2,\cdots)$，零点能为 $\dfrac{1}{2}\hbar\omega$。于是这个盒子看成充满了能量为 $n\hbar\omega$ 的光子。光子与物质发生相互作用时，发射或吸收一个光子，使第 n 类光子的数目改变 ±1。光子与物质相互作用的 Hamilton 量可用微扰的办法得到。

为方便计算量子电动力学基本过程，Feynman 发明了 Feynman 图，他取 x 轴或 y 轴代表时间，用带箭头的实线表示电子或正电子，如果箭头指向与时间轴正方向一致，则该实线代表运动的电子，如果相反则表示运动的正电子，Feynman 图中波浪线表示光子，封闭圈表示正负电子对，如图 D.1 所示。

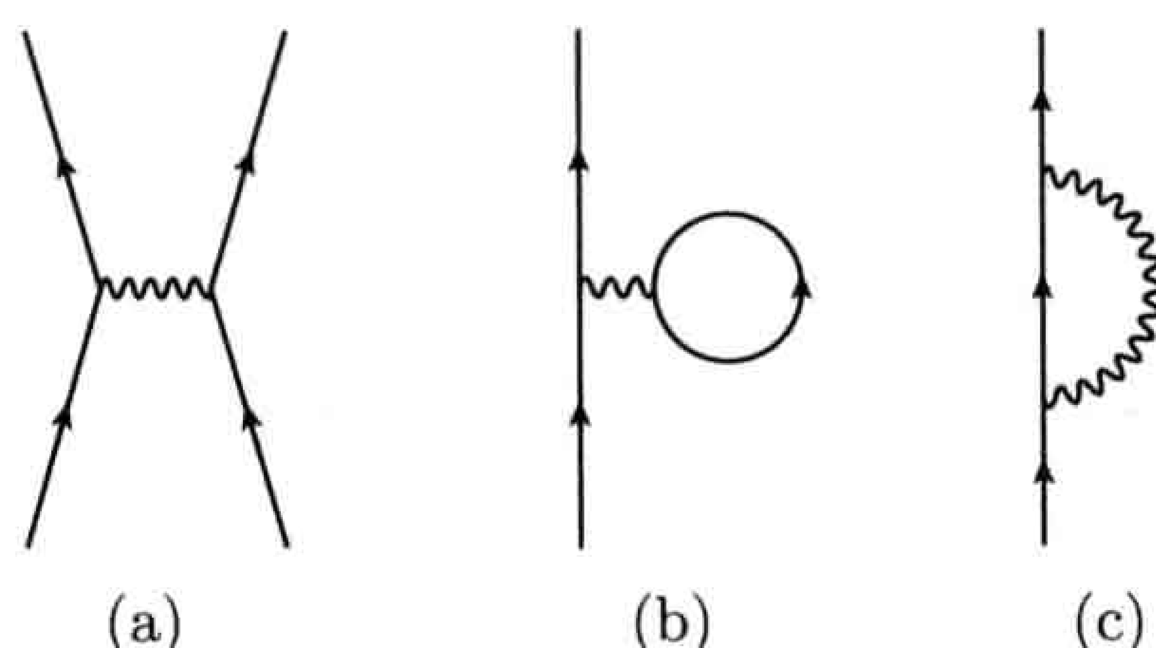


图 D.1　单个光子和电子的相互作用 Feynman 图

图 D.1 描述了单个光子和电子的相互作用，其中竖直轴为时间轴，双光子与电子作用的图过于复杂，在此省略。图 D.1 中 (a) 表示两个电子通过交换一个光子发生相互作用；(b) 显示的是电子与场基态的相互作用。场基态又定义为真空，与我们平时理解的真空必空无一物不同，量子电动力学的真空中包含存活时间很短的正负电子对，从真空中产生并彼此湮灭。图 (b) 中正负电子对湮灭时产生一个光子，而光子又电子散射 —— 此过程称为真空极化；(c) 代表电子发射然后再吸收同一光子，称为电子的自能。

在第 3 章我们给出了图 (a) 的计算公式 (见 (3.50) 式)，并由此得到了精确到 $1/c^2$ 量级的 Dirac-Coulomb-Breit (DCB) Hamilton 算符 —— 也是目前用于分子计

① 详见 Greiner 和 Reinhardt 所著的 *Quantum Electrodynamics*. Berlin: Springer-Verlag, 1992

算的最精确 Hamilton 算符。真空极化 (见图 (b)) 和电子的自能 (见图 (c)) 是相对论效应的高阶校正，我们称为 QED 效应。QED 效应的计算十分复杂，从化学精度考量，它们均为小量，在相对论量子化学计算中可以忽略不计。下面我们通过计算氢原子超精细光谱中的 Lamb 移位 —— 即氢原子的 $2s_{1/2}$ 与 $2p_{1/2}$ 的能级差，来说明 QED 理论的精确性。

对轻原子 ($Z\alpha \ll 1$)，采用一级微扰理论，真空极化对原子能级的影响可用以下公式表示

$$\Delta E_{nl}^{\mathrm{VP}} = -\frac{4m_{\mathrm{e}}}{15\pi n^3}\alpha(Z\alpha)^4\delta_{l0} \tag{D.1}$$

式中，α 为精细结构常数；n, l 分别为轨道主量子数和角量子数。由于 δ_{l0} 函数仅在 $l = 0$ 时不为零，所以真空极化只影响 s 轨道的能级。注意上式采用的是自然单位制①，考虑氢原子的 $2s_{1/2}$ 轨道 ($Z = 1, n = 2, l = 0$)，代入 (D.1) 式得

$$\Delta E_{2s_{1/2}}^{\mathrm{VP}} = -1.122 \times 10^{-7}\ \mathrm{eV} = -27.1\ \mathrm{MHz} \tag{D.2}$$

上式第二个等号利用 $E = 2\pi\hbar\omega$ 将能量 (eV) 转为频率 (MHz)。由于 $2p_{1/2}$ 轨道能级不受真空极化的影响，所以 (D.2) 式即为真空极化对氢原子 Lamb 移位的贡献。

如果将自然单位制换为原子单位制，真空极化对原子能级的贡献 (D.1) 式变为

$$\begin{aligned}\Delta E_{nl}^{\mathrm{VP}} &= -\frac{4c^2}{15\pi n^3}\alpha(Z\alpha)^4\delta_{l0} \\ &= -\frac{4}{15\pi n^3}Z^4\alpha^3\delta_{l0}\end{aligned} \tag{D.3}$$

由上式可见，真空极化对氢原子能级的影响在 α^3 数量级。

电子自能对原子能级的影响可表示为 (自然单位制)

$$\begin{aligned}\Delta E_{njl}^{\mathrm{SE}} = \frac{4m_{\mathrm{e}}}{3\pi n^3}\alpha(Z\alpha)^4\Big\{L_{nl} &+ \left[\frac{5}{6} - 2\ln(Z\alpha)\right]\delta_{l0} \\ &\pm \frac{3}{4}\frac{1}{(2j+1)(2l+1)}(1-\delta_{l0})\Big\}\end{aligned} \tag{D.4}$$

式中，$j = l \pm 1/2$

$$\begin{aligned}L_{nl} = \frac{n^3}{2m_{\mathrm{e}}^2(Z\alpha)^4}&\sum_{n'l'}\left|\langle \Psi'_{n'l'}|\hat{\boldsymbol{p}}|\Psi_{nl}\rangle\right|^2 \\ &\times (E_{n'l'} - E_{nl})\ln\frac{(Z\alpha)^2 m_{\mathrm{e}}}{2(E_{n'l'} - E_{nl})}\end{aligned} \tag{D.5}$$

此式也称为 Bethe 对数，只能用数值法求解。与真空极化类似，电子自能对原子能级的影响也在 α^3 数量级。

① 在自然单位制中，$c = \hbar = 1$；$m_{\mathrm{e}} = 5.117 \times 10^5$；能量单位为电子伏特 (eV)

对氢原子的 $2s_{1/2}$ 和 $2p_{1/2}$ 轨道，数值法求得

$$\begin{aligned} L_{20} &\approx -2.811\ 77 \\ L_{21} &\approx 0.030\ 02 \end{aligned} \tag{D.6}$$

代入 (D.4) 式，得电子自能对 $2s_{1/2}$ 和 $2p_{1/2}$ 的能级影响分别为

$$\begin{aligned} \Delta E_{2s_{1/2}} &= 5.420 \times 10^{-6}\ \text{eV} = 1\ 066.4\ \text{MHz} \\ \Delta E_{2p_{1/2}} &= -5.328 \times 10^{-8}\ \text{eV} = -12.9\ \text{MHz} \end{aligned} \tag{D.7}$$

二者相减，得电子自能对 Lamb 移位的贡献为 1 079.3 MHz，再加上真空极化对 Lamb 移位的贡献 (见 (D.2) 式)，得 Lamb 移位计算值为 1 052.2 MHz。如果考虑更高阶的校正 (到 α^6 数量级)，计算值可修正为

$$L^{\text{theor}} = 1\ 057.864 \pm 0.014\ \text{MHz} \tag{D.8}$$

与实验值 1 057.8 MHz 几乎完全一样。Lamb 移位与精细结构分裂 $\Delta\omega(2p_{3/2} - 2p_{1/2}) = 10\ 950$ MHz 相比，小一个量级。

对超重元素，QED 效应逐渐增大，关于此方面研究的进展读者可参阅文献①。

① Indelicato P, Santos J P, Boucard S, et al. QED and relativistic corrections in superheavy elements. Eur Phys J D, 2007, 45: 155~170

附录 E　Hylleraas-Undheim-MacDonald 定理

E.1　线性变分法

假设体系的 Hamilton 算符为 $\hat{H}$，其归一化本征函数 $\varPsi$ 可写成一组正交归一化的函数 $\{\varPhi_n\}(n=1,2,\cdots,m)$ 的线性组合

$$\varPsi=\sum_{n=1}^{m}c_n\varPhi_n \tag{E.1}$$

式中，c_n 为待定实参数，且 $\sum\limits_{n=1}^{m}c_n^2=1$。由能量最低原理：用任何近似状态波函数计算的体系能量一定大于或等于真正的基态本征态的能量，即

$$W=\sum_i\sum_j c_ic_jH_{ij}\geqslant E_0 \tag{E.2}$$

式中

$$H_{ij}=\langle\varPhi_i|\hat{H}|\varPhi_j\rangle \tag{E.3}$$

可改变上式中的 c_i,c_j 求得 W 的最小值，这样的变分法称为线性变分法。
要使 W 最小，必须使 $\dfrac{\partial W}{\partial c_j}=0(j=1,2,\cdots,m)$，即

$$\sum_i\left(H_{ij}-W\delta_{ij}\right)c_i=0\qquad(i,j=1,2,\cdots,m) \tag{E.4}$$

式中，δ_{ij} 函数的取值为

$$\delta_{ij}=\begin{cases}1 & (i=j)\\ 0 & (i\neq j)\end{cases} \tag{E.5}$$

这是含有 m 个独立变量的 $c_1,c_2,\cdots,c_m$ 齐次线性联立方程组。如果要使这个方程

组有非零解，则本征行列式必须为零，即

$$|H_{ij}-W| \equiv \begin{vmatrix} H_{11}-W & H_{12} & \cdots & H_{1m} \\ H_{21} & H_{22}-W & \cdots & H_{2m} \\ \vdots & \vdots & & \vdots \\ H_{m1} & H_{m2}-W & \cdots & H_{mm}-W \end{vmatrix} = 0 \tag{E.6}$$

(E.6) 式称为本征方程 (或久期方程)，求解此方程将得到 m 个实根，将这 m 个实根按大小顺序排列，得

$$W_{m,1} \leqslant W_{m,2} \leqslant W_{m,3} \leqslant \cdots \leqslant W_{m,m} \tag{E.7}$$

m 个实根的最小值 $W_{m,1}$ 是体系基态能量 E_0 的上限。相应于 m 个实根的正交归一化的本征函数记为 $\Psi_{m1}, \Psi_{m\ 2}, \cdots, \Psi_{mm}$。

E.2 Hylleraas-Undheim-MacDonald 定理的推导

考虑在 m 个正交归一化的函数 $\{\Phi_n\}(n=1,2,\cdots,m)$ 中再增加一个函数 Φ_{m+1}，并假定 Φ_{m+1} 与其他 $\{\Phi_n\}(n=1,2,\cdots,m)$ 也正交且归一化。为讨论方便，现在用 $m+1$ 个正交归一化函数 $\Phi_{m+1}, \{\Psi_{mn}\}(n=1,2,\cdots,m)$ 来展开 Ψ，利用线性变分法，可得到与 (E.6) 式类似的本征方程

$$\begin{aligned} D_{m+1}(W) &= \begin{vmatrix} W_{m,1}-W & 0 & \cdots & 0 & h_1 \\ 0 & W_{m,2}-W & \cdots & 0 & h_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & W_{m,m}-W & h_m \\ h_1* & H_2* & \cdots & h_m* & h_{m+1}-W \end{vmatrix} \\ &= G_{m+1}(W) - F_{m-1}(W) = 0 \end{aligned} \tag{E.8}$$

式中

$$\begin{aligned} h_r &= \langle \Psi_{m,r}|\hat{H}|\Phi_{m+1}\rangle \qquad (r=1,2,\cdots,m) \\ h_{m+1} &= \langle \Phi_{m+1}|\hat{H}|\Phi_{m+1}\rangle \\ G_{m+1}(W) &= (h_{m+1}-W)\prod_{k=1}^{m}(W_{m,k}-W) \\ F_{m-1}(W) &= \sum_{r=1}^{m}|h_r|^2\prod_{k\neq r}^{m}(W_{m,k}-W) \end{aligned} \tag{E.9}$$

求解本征方程 (E.8) 式可得到 $m+1$ 个实根，这 $m+1$ 个实根按大小顺序排列，得

$$W_{m+1,1} \leqslant W_{m+1,2} \leqslant W_{m+1,3} \leqslant \cdots \leqslant W_{m+1,m+1} \tag{E.10}$$

为比较这 $m+1$ 个实根与求解 (E.6) 式得到的 m 个实根的关系，考虑当 W 取不同值时行列式 $D_{m+1}(W)$ 的值：

(1) 如果 $W = W_{m,r}$，得

$$D_{m+1}(W_{m,r}) = -F_{m-1}(W_{m,r}) = -|h_r|^2 \prod_{k\neq r}^{m} (W_{m,k} - W_{m,r}) \tag{E.11}$$

下面讨论 $D_{m+1}(W_{m,r})$ 的正负号。由于 m 个实根是顺序排列的，当 $k < r$ 时，上式连乘号内的 $(W_{m,k} - W_{m,r}) \leqslant 0$，当 $k > r$ 时，$(W_{m,k} - W_{m,r}) \geqslant 0$。所以当 r 为奇数时，连乘号内小于零的项有偶数个，最后乘积的结果大于或等于零，即 $D_{m+1}(W_{m,r}) \geqslant 0$；当 r 为偶数时，连乘号内小于零的项有奇数个，最后乘积的结果小于或等于零，即 $D_{m+1}(W_{m,r}) \leqslant 0$。由此可以推测，$D_{m+1}(W) = 0$ 的根 $W_{m,r} \leqslant W \leqslant W_{m,r+1}$。

(2) 如果 $W \to -\infty$，则 $D_{m+1}(W) \to \infty$。

(3) 如果 $W \to \infty$，则 $D_{m+1}(W) \to (-1)^{m+1}\infty$。

综合以上三条，再考虑到 $D_{m+1}(W) = 0$ 只有 $m+1$ 个实根，我们可以肯定这 $m+1$ 个实根必穿插在数列 $-\infty, W_{m,1}, W_{m,2}, \cdots, W_{m,m}, \infty$ 的两个相邻数之间，即

$$W_{m+1,1} \leqslant W_{m,1} \leqslant W_{m+1,2} \leqslant W_{m,2} \leqslant \cdots \leqslant W_{m+1,m} \leqslant W_{m,m} \leqslant W_{m+1,m+1} \tag{E.12}$$

上式即为 Hylleraas-Undheim-MacDonald (HUM) 定理。由 (E.12) 式我们可以得出重要推论：$W_{m,1}$ 是基态能量 E_0 的上限，其他根 $W_{m,2}, \cdots, W_{m,m}$ 分别为正确激发态的本征能级 $E_1, E_2, \cdots, E_{m-1}$ 的上限。相应的本征函数 $\Psi_{m,2}, \Psi_{m,3}, \cdots, \Psi_{m,m}$ 可作为各激发态的近似波函数。

附录 F　Lagrange 方法

Lagrange 方法又称为 Lagrange 乘子法，以二元函数 $f(x,y)=xy$ 为例，欲求二元函数 $f(x,y)$ 在约束条件 $g(x,y)=x^2+y^2=1$ 下的极值，即找出 x,y 分别取什么值时 $f(x,y)$ 达到最大 (或最小) 值。

图 F.1 画出了二元函数 $f(x,y)$ 的等高线图 (实线，以 $d_n(n=1\sim6)$ 标记)，图中不同 $d_n(n=1\sim6)$ 分别对应于

$$\begin{aligned} f(x,y) &= f_n \quad (n=1\sim6) \\ &= -1,-0.5,-0.2,0.2,0.5,1 \end{aligned} \tag{F.1}$$

图 F.1 中还画出了 $g(x,y)=1$ 的等高线，用虚线表示。当虚线与 d_n 相交时，表示方程组 $f(x,y)=f_n$，$g(x,y)=1$ 有解。由于 $f(x,y)$ 是连续函数，与虚线相交的等高线不只一条 —— 显然，图 F.1 中的 d_2,d_3,d_4,d_5 等均和虚线相交。只有当 $f(x,y)$ 的等高线与 $g(x,y)=1$ 相切时，$f(x,y)$ 才出现极值，从图 F.1 可以看出，等高线 d_2 和 d_5 均与虚线相切，它们对应的值 -0.5, 0.5 分别为满足约束条件的 $f(x,y)$ 的极小和极大值。

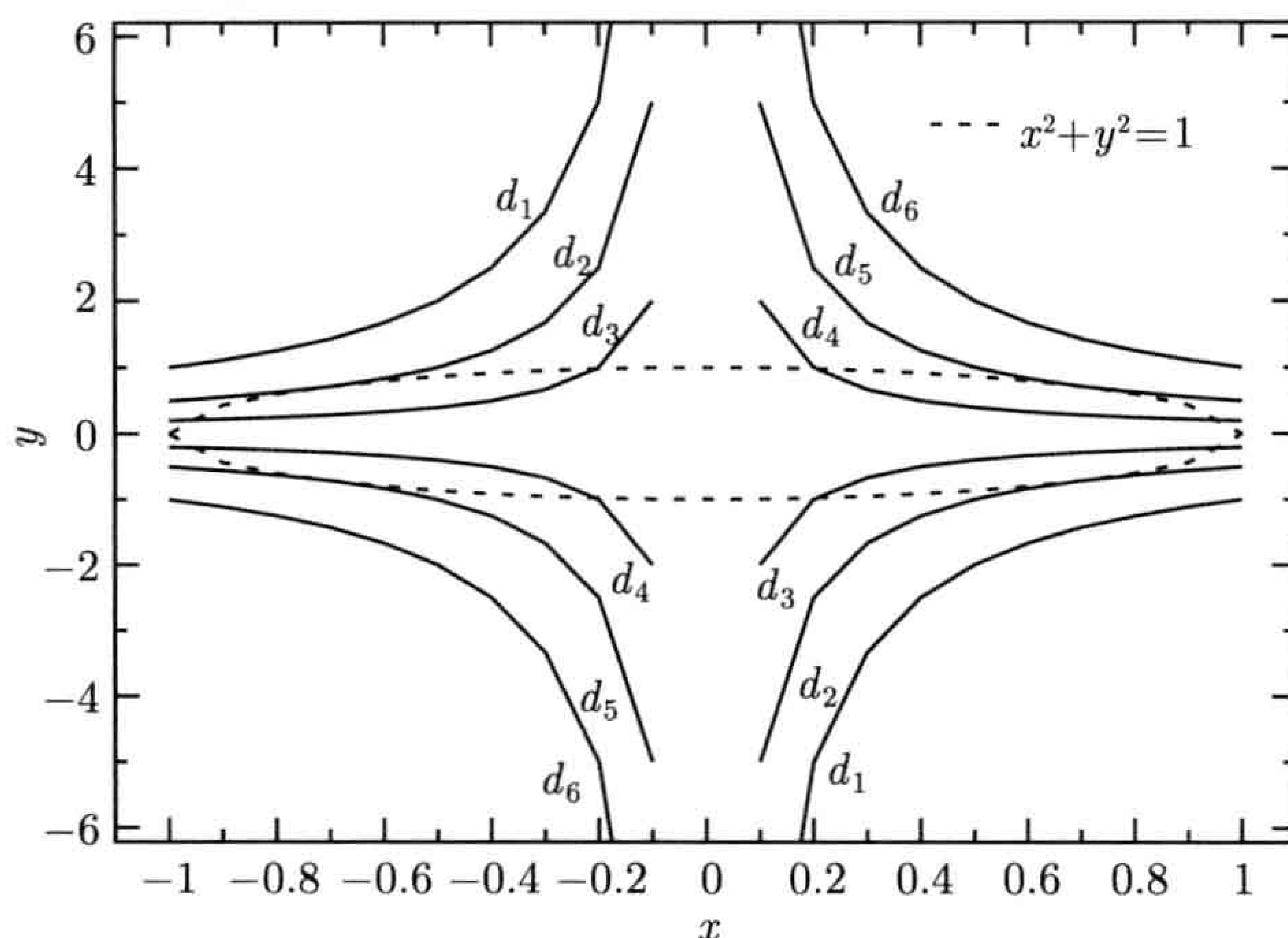


图 F.1　二元函数 $f(x,y)$ 的等高线图

用矢量的形式来表达的话，$f(x,y)$ 的等高线与 $g(x,y)=c$ 在某点 P 相切意味着两条曲线的斜率在 P 点平行。此时引入 Lagrange 乘子 λ, $(\lambda\neq0)$，并求解

$$\nabla L(x,y,\lambda)\equiv\nabla\left[f(x,y)+\lambda\left(g(x,y)-c\right)\right]=0 \tag{F.2}$$

求出 λ 及极值对应的点 (x, y) 后，将其代入 $f(x, y)$，即可求出极值。

再回到我们的例子，用 Lagrange 乘子法可求得

$$\lambda = \pm 1/2, \quad x = \pm\sqrt{2}/2, \quad y = \pm\sqrt{2}/2 \tag{F.3}$$

易知函数的最小值为 -0.5，最大值为 0.5，和图 F.1 的结论一致。

推广到一般情形，对多元函数 $f(x_1, x_2, \cdots, x_n)$，约束条件有 M 个，用 $g_k(x_1, x_2, \cdots, x_n) = c_k (k = 1, 2, \cdots, M)$ 表示，试求 $f(x_1, x_2, \cdots, x_n)$ 的极值。

引入 M 个 Lagrange 乘子 $\lambda_k (k = 1, 2, \cdots, M)$，得

$$\begin{aligned} L(x_1, x_2, \cdots, x_n, \lambda_1, \lambda_2, \cdots, \lambda_M) = & f(x_1, x_2, \cdots, x_n) \\ & + \sum_{k=1}^{M} \lambda_k \left[g_k(x_1, x_2, \cdots, x_n) - c_k \right] \end{aligned} \tag{F.4}$$

为得到极值条件，我们需求解

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= 0 \quad (i = 1, 2, \cdots, n) \\ \frac{\partial L}{\partial \lambda_k} &= 0 \quad (k = 1, 2, \cdots, m) \end{aligned} \tag{F.5}$$

这是一个含有 $(n+m)$ 个未知数的 $(n+m)$ 个方程的方程组，求出 $\lambda_k (k = 1, 2, \cdots, M)$ 及 $x_i (i = 1, 2, \cdots, n)$ 后，代入 $f(x_1, x_2, \cdots, x_n)$，可求出极值。

附录 G　动 量 表 象

通俗的说，位置空间是由 (x, y, z) 作为坐标的空间，动量空间则是用动量的分量 (p_x, p_y, p_z) 作为坐标的空间。动量空间中的一个点代表了一个动量，是一种直观的表示方法。对每一个算符 $\hat{O}$ (或态函数 $\varPhi$)，我们称 $\hat{O}(\boldsymbol{r})$(或 $\varPhi(\boldsymbol{r})$) 是算符 (或态函数) 在坐标表象 ($\boldsymbol{r}$ 表象) 中的表示，而 $\hat{O}(\boldsymbol{p})$(或 $\varPhi(\boldsymbol{p})$) 是算符 (或态函数) 在动量表象 ($\boldsymbol{p}$ 表象) 中的表示。读者一般熟悉前者，下面我们就从算符 (或态函数) 在 $\boldsymbol{r}$ 表象中的表示出发来得到它们在 $\boldsymbol{p}$ 表象中的表示。为简化讨论，我们这里仅考虑不含时间的算符和函数。

G.1　态　函　数

在 $\boldsymbol{r}$ 表象中动量算符 $\hat{\boldsymbol{p}}$ 的本征函数 $\varPsi_p(\boldsymbol{r})$ 为平面波形式 (见第 2 章 2.3.1 节)

$$\varPsi_p(\boldsymbol{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} \tag{G.1}$$

由量子力学的态叠加原理，任何一个态函数 $\varPhi$ 都可以看成是不同动量的平面波叠加。换句话说，任何波函数 $\varPhi$ 都可以写成如下形式

$$\varPhi(\boldsymbol{r}) = \int \varPhi(\boldsymbol{p})\, \varPsi_p(\boldsymbol{r}) \mathrm{d}\boldsymbol{p} \tag{G.2}$$

式中，$\varPhi(\boldsymbol{p})$ 即为 $\varPhi$ 在动量空间中的表示，可由下式得出

$$\varPhi(\boldsymbol{p}) = \int \varPhi(\boldsymbol{r})\, \varPsi_p^*(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \tag{G.3}$$

由 (G.2) 式和 (G.3) 式可见，$\varPhi$ 在 $\boldsymbol{r}$ 表象和 $\boldsymbol{p}$ 表象中的表示 $\varPhi(\boldsymbol{r})$, $\varPhi(\boldsymbol{p})$ 互为 Fourier 变换。

G.2　算　　符

在 $\boldsymbol{r}$ 表象中，将算符 $\hat{O}$ 作用到态函数 $\varPhi$，得

$$\hat{O}(\boldsymbol{r})\, \varPhi(\boldsymbol{r}) = \varPhi'(\boldsymbol{r}) \tag{G.4}$$

方程 (G.4) 式在 $\boldsymbol{p}$ 表象中的表示可如下求出。$\Phi(\boldsymbol{r})$, $\Phi'(\boldsymbol{r})$ 作 Fourier 变换

$$\begin{aligned}\Phi(\boldsymbol{r}) &= \frac{1}{(2\pi\hbar)^{3/2}}\int \Phi(\boldsymbol{p}')\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}'\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{p}' \\ \Phi'(\boldsymbol{r}) &= \frac{1}{(2\pi\hbar)^{3/2}}\int \Phi'(\boldsymbol{p}')\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}'\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{p}'\end{aligned} \tag{G.5}$$

代入 (G.4) 式，得

$$\frac{1}{(2\pi\hbar)^{3/2}}\int \mathrm{d}\boldsymbol{p}'\hat{O}(\boldsymbol{r})\Phi(\boldsymbol{p}')\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}'\cdot\boldsymbol{r}} = \frac{1}{(2\pi\hbar)^{3/2}}\int \mathrm{d}\boldsymbol{p}'\,\Phi'(\boldsymbol{p}')\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}'\cdot\boldsymbol{r}} \tag{G.6}$$

两边乘以 $\mathrm{e}^{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}$，再对 $\boldsymbol{r}$ 积分，得

$$\int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{p}'\hat{O}(\boldsymbol{r})\Phi(\boldsymbol{p}')\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} = \int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{p}'\,\Phi'(\boldsymbol{p}')\mathrm{e}^{-\frac{i}{\hbar}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} \tag{G.7}$$

利用 Dirac δ 函数的性质

$$\int \mathrm{d}\boldsymbol{r}\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} = (2\pi\hbar)^3\delta(\boldsymbol{p}-\boldsymbol{p}') \tag{G.8}$$

代入 (G.7) 式，整理得

$$\int \mathrm{d}\boldsymbol{p}'\hat{O}(\boldsymbol{p},\boldsymbol{p}')\Phi(\boldsymbol{p}') = \Phi'(\boldsymbol{p}) \tag{G.9}$$

由上可知，动量表象中的算符 $\hat{O}$ 可表示为积分算符形式，式中 $\hat{O}(\boldsymbol{p},\boldsymbol{p}')$ 称为 $\hat{O}$ 的核，定义为

$$\hat{O}(\boldsymbol{p},\boldsymbol{p}') = \frac{1}{(2\pi\hbar)^{3/2}}\int \mathrm{d}\boldsymbol{r}\hat{O}(\boldsymbol{r})\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} \tag{G.10}$$

下面讨论几个常见算符在动量表象中的表示。

1. **动量算符**

将 $\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla$ 代入 (G.10) 式，并利用 (G.8) 式，得动量算符 $\hat{\boldsymbol{p}}$ 在动量空间的核为

$$\begin{aligned}\hat{\boldsymbol{p}}(\boldsymbol{p},\boldsymbol{p}') &= \frac{-\mathrm{i}}{(2\pi\hbar)^{3/2}}\int \mathrm{d}\boldsymbol{r}\nabla\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} \\ &= \delta(\boldsymbol{p}-\boldsymbol{p}')\end{aligned} \tag{G.11}$$

再代入 (G.9) 式，得

$$\int \mathrm{d}\boldsymbol{p}'\hat{\boldsymbol{p}}(\boldsymbol{p},\boldsymbol{p}')\Phi(\boldsymbol{p}') = \boldsymbol{p}\Phi(\boldsymbol{p}) = \Phi'(\boldsymbol{p}) \tag{G.12}$$

即在动量表象中动量算符可直接表示为动量 $\boldsymbol{p}$。

2. 位置算符

将 $\hat{\boldsymbol{r}}$ 代入 (G.10) 式，再代入 (G.9) 式，得

$$\begin{aligned}\int \mathrm{d}\boldsymbol{p}'\hat{\boldsymbol{r}}(\boldsymbol{p},\boldsymbol{p}')\varPhi(\boldsymbol{p}') &= \int \mathrm{d}\boldsymbol{p}'\frac{1}{(2\pi\hbar)^{3/2}}\int \mathrm{d}\boldsymbol{r}\boldsymbol{r}\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}}\varPhi(\boldsymbol{p}') \\ &= \mathrm{i}\hbar\frac{\partial}{\partial \boldsymbol{p}}\varPhi(\boldsymbol{p}) = \varPhi'(\boldsymbol{p})\end{aligned} \tag{G.13}$$

上式再次利用了 (G.8) 式。由 (G.13) 式知，在动量表象中位置算符表示为微分算符 $\mathrm{i}\hbar\frac{\partial}{\partial \boldsymbol{p}}$。

3. Coulomb 势

为简洁起见，取 $\hbar=1$(原子单位)。在 $\boldsymbol{r}$ 表象中 Coulomb 势 $V(\boldsymbol{r})=-Z/r$(原子单位)，代入 (G.10) 式得动量表象中 V 的核为

$$V(\boldsymbol{p},\boldsymbol{p}') = \frac{1}{(2\pi)^{3/2}}\int \mathrm{d}\boldsymbol{r}\,(-Z/r)\,\mathrm{e}^{-\mathrm{i}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} \tag{G.14}$$

为求得上述积分，引入切断函数 $\mathrm{e}^{-kr}, k>0$。令 $V'(\boldsymbol{r})=-Z\mathrm{e}^{-kr}/r$，显然，当 $k\to 0$ 时，$V'(\boldsymbol{r})\to V(\boldsymbol{r})$。动量表象中 V' 的核为

$$\begin{aligned}V'(\boldsymbol{p},\boldsymbol{p}') &= -\frac{1}{(2\pi)^{3/2}}\int \mathrm{d}\boldsymbol{r}\frac{Z\mathrm{e}^{-kr}}{r}\mathrm{e}^{-\mathrm{i}(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} \\ &= -\frac{2Z\pi}{(2\pi)^{3/2}}\int_0^\infty r\mathrm{d}r\mathrm{e}^{-kr}\int_0^\pi \mathrm{e}^{-\mathrm{i}\left(|\boldsymbol{p}-\boldsymbol{p}'|r\cos\theta\right)}\sin\theta\mathrm{d}\theta\end{aligned} \tag{G.15}$$

上式中

$$\int_0^\pi \mathrm{e}^{-\mathrm{i}\left(|\boldsymbol{p}-\boldsymbol{p}'|r\cos\theta\right)}\sin\theta\mathrm{d}\theta = \frac{-1}{\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|r}\left(\mathrm{e}^{-\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|r}-\mathrm{e}^{\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|r}\right) \tag{G.16}$$

代入 (G.15) 式，得

$$\begin{aligned}V'(\boldsymbol{p},\boldsymbol{p}') &= \frac{Z2\pi}{(2\pi)^{3/2}\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|}\int_0^\infty \mathrm{d}r\mathrm{e}^{-kr}\left(\mathrm{e}^{-\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|r}-\mathrm{e}^{\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|r}\right) \\ &= -Z\mathrm{i}\left(\frac{1}{2\pi}\right)^{1/2}\cdot\frac{1}{|\boldsymbol{p}-\boldsymbol{p}'|}\int_0^\infty \mathrm{d}r\left(\mathrm{e}^{-(k+\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|)r}-\mathrm{e}^{-(k-\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|)r}\right) \\ &= -Z\mathrm{i}\left(\frac{1}{2\pi}\right)^{1/2}\cdot\frac{1}{|\boldsymbol{p}-\boldsymbol{p}'|}\left(\frac{1}{k+\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|}-\frac{1}{k-\mathrm{i}|\boldsymbol{p}-\boldsymbol{p}'|}\right) \\ &= -Z\left(\frac{2}{\pi}\right)^{1/2}\cdot\frac{1}{k^2+|\boldsymbol{p}-\boldsymbol{p}'|^2}\end{aligned} \tag{G.17}$$

令上式中 $k=0$，得动量表象中 Coulomb 势 $V(\boldsymbol{r})=-Z/r$ 的核为

$$V(\boldsymbol{p},\boldsymbol{p}') = -Z\left(\frac{2}{\pi}\right)^{1/2}\cdot\frac{1}{|\boldsymbol{p}-\boldsymbol{p}'|^2} \tag{G.18}$$

索　　引

F

G

H

J

K

L

Z

其他